



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 171531

TO: Ginny Portner
Location: rem/3B02/3C18
Art Unit: 1645
Friday, December 02, 2005
Case Serial Number: 09/904994

From: Toby Port
Location: Biotech-Chem Library
REM-1A59
Phone: 571-272-2523

toby.port@uspto.gov

Search Notes

Examiner Portner,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Toby Port
X22523



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact*:

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library, Remsen Bldg.



STIC-Biotech/ChemLib

171531

MS

From: STIC-ILL
Sent: Monday, November 14, 2005 5:36 PM
To: STIC-Biotech/ChemLib
Subject: FW: 09/904,994

RECEIVED
NOV 15 2005
STIC-BIOTECH/CHEM LIB
(STIC)

-----Original Message-----

From: Portner, Ginny
Sent: Monday, November 14, 2005 5:29 PM
To: STIC-ILL
Subject: 09/904,994

Please search SEQ ID NO 1 with respect to Nucleic acid claims, and fragments thereof. thanks

Ginny Portner
Remsen Building
Art Unit 1645
Room E03, B02
(571) 272-0862

3C18

11/15/2005

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 23:29:19 ; Search time 14055 Seconds
(without alignments)
11659.884 Million cell updates/sec

Title: US-09-904-994B-1

Perfect score: 2883

Sequence: 1 rgragatttccarcactt.....aaaaagtagagccacagg 2883

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_scs.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_hcg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2880.6	99.9	2883	6	BD185302	Helicobac
2	2880.6	99.9	2883	6	AX356683	Sequence
3	2228	77.3	2452	6	BD185306	Helicobac
4	2228	77.3	2452	6	AX356695	Sequence
5	2158.6	74.9	2405	6	BD185303	Helicobac
6	2158.6	74.9	2405	6	AX356686	Sequence
7	2158.6	74.9	2407	6	BD185305	Helicobac
8	2158.6	74.9	2407	6	AX356692	Sequence
9	1948	67.6	2183	6	BD185304	Helicobac
10	1948	67.6	2183	6	AX356689	Sequence
11	1185.6	41.1	8406	1	AF330621	Helicobac
12	1138	39.5	2664	1	HECUREASE	Helicobac
13	1090.8	37.8	12037	1	AE001446	Helicobac
14	1081.6	37.5	10903	1	AE000529	Helicobac
15	1075.2	37.3	2735	6	AR054309	Sequence
16	1075.2	37.3	5100	1	HECUREABCD	Helicobac
17	1072.4	37.2	7088	1	AB032429	Helicobac
18	1062.6	36.9	2619	6	A41006	Sequence 1

19	1062.6	36.9	2619	6	A72509	Sequence 1
20	1062.6	36.9	2619	6	A73285	Sequence 1
21	1062.6	36.9	2619	6	AR061045	Sequence
22	1062.6	36.9	2619	6	AR161678	Sequence
23	1060.4	36.8	2430	1	AY227442	Helicobac
24	1058.6	36.7	4824	6	AR349305	Sequence
25	1058.6	36.7	4824	6	AR349306	Sequence
26	1057.8	36.7	2619	1	HFHURE	Sequence
27	1056.8	36.7	2746	1	HPHPUABG	Sequence
28	1056.8	36.7	2767	6	A31515	Sequence
29	985.2	34.2	6126	1	AF332656	Helicobac
30	985.2	34.2	308929	1	AE017145	Helicobac
31	894.6	31.0	1704	1	AY714224	Helicobac
32	885.8	30.7	1815	6	BD092530	Identific
33	884.8	30.7	1710	1	AY295085	Helicobac
34	881.8	30.6	1707	6	CQ903244	Sequence
35	881.8	30.6	1707	6	CQ903245	Sequence
36	880	30.5	1710	1	AF352376	Helicobac
37	878.4	30.5	1710	6	A07398	Sequence
38	878.4	30.5	1710	6	A08818	Sequence
39	876	30.4	1680	1	AF479026	Helicobac
40	772.4	26.8	110000	1	BA000004_02	Continuation (3 of
41	766	26.6	10118	1	BSUPEROP	Sequence
42	766	26.6	200690	1	BSUB0019	Sequence
43	739.4	25.6	1665	1	AF508003	Helicobac
44	737.6	25.6	6537	1	BACUREA	Sequence
45	733.8	25.5	4723	1	AB201709	Sequence

ALIGNMENTS

RESULT 1	BD185302	BD185302	2883 bp	DNA	linear	PAT 17-JUN-2003
LOCUS	BD185302	Helicobacter felis vaccine.				
DEFINITION	BD185302	Helicobacter felis vaccine.				
ACCESSION	BD185302	Helicobacter felis vaccine.				
VERSION	BD185302.1	GI:31877502				
KEYWORDS	JP 2002355054-A/1.					
SOURCE	JP 2002355054-A/1.					
ORGANISM	Helicobacter felis					
REFERENCE	1 (bases 1 to 2883)					
AUTHORS	Kusters,J.G. and Cattoli,G.					
TITLE	Helicobacter felis vaccine					
JOURNAL	Patent: JP 2002355054-A 1 10-DEC-2002;					
COMMENT	AKZO NOBEL NV					
	OS Helicobacter felis					
	PN JP 2002355054-A/1					
	PD 10-DEC-2002					
	PF 16-JUL-2001 JP 2001214711					
	PR 17-JUL-2000 EP 00202565.8					
	PT JOHANNES GERARDUS KUSTERS,GIOVANNI CATTOLI					
	PC					
	C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC					
	A61K39/23,A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15, PC					
	C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/53, PC					
	G01N33/53,G01N33/566,G01N33/569/(C12N9/80,C12R1/01), (C12Q1/68, PC					
	C12R1/01),					
	PC C12N15/00,C12N5/00,A61K37/02					
	CC Helicobacter felis vaccine					
	FT Key					
	FT CDS					
	FT CDS					
	Location/Qualifiers					
	1..2883					
	/organism="Helicobacter felis"					

/mol_type="genomic DNA" /db_xref="taxon:214"									
Query Match 99.9%; Score 2880.6; DB 6; Length 2883; Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 2883; Conservative 0;									
QY	1	RGRAGATTTTCCARCACTTCAAGCACATATTCGTCGTGTTGGTGGTGGTAAATTCR	60						
DB	1	RGRAGATTTTCCARCACTTCAAGCACATATTCGTCGTGTTGGTGGTAAATTCR	60						
QY	61	ACTTGTAATRCATTAATTAATTTTAAATAATTAATTAATTAATTAATTAATTA	120						
DB	61	ACTTGTAATRCATTAATTAATTTTAAATAATTAATTAATTAATTAATTAATTA	120						
QY	121	TTACTTTATTAATAAGATTAATAAGTAACGAAATTAAGACTATAATCCCATTCGCTT	180						
DB	121	TTACTTTATTAATAAGATTAATAAGTAACGAAATTAAGACTATAATCCCATTCGCTT	180						
QY	181	TAAAAATTTAACACAAGGAGTAATAGGTGAAACTCACACCCAAAGAGCAAGAAAAGTTCTT	240						
DB	181	TAAAAATTTAACACAAGGAGTAATAGGTGAAACTCACACCCAAAGAGCAAGAAAAGTTCTT	240						
QY	241	GTATATATATGCGGGCGAAGTGCTGAGAAAGCGGAAAGCAGAGGCTTAAAGCTCAACCA	300						
DB	241	GTATATATATGCGGGCGAAGTGCTGAGAAAGCGGAAAGCAGAGGCTTAAAGCTCAACCA	300						
QY	301	ACCGAAGCATTTGCTTACATTAATAGTCCCATATATATGGAAGAGCGCGCGTGGAAAAA	360						
DB	301	ACCGAAGCATTTGCTTACATTAATAGTCCCATATATATGGAAGAGCGCGCGTGGAAAAA	360						
QY	361	AACCGTTGCCAGCTTATGGAAGTGTCATGCACTTTTGAAAAAGATGAAGTAATGCC	420						
DB	361	AACCGTTGCCAGCTTATGGAAGTGTCATGCACTTTTGAAAAAGATGAAGTAATGCC	420						
QY	421	CGGGTGGTAAATATGGTTCCTCCGATCTAGGTGTAAGCCACTTTCTGATGTACGAA	480						
DB	421	CGGGTGGTAAATATGGTTCCTCCGATCTAGGTGTAAGCCACTTTCTGATGTACGAA	480						
QY	481	ACTTGTAACGTGTAATGGGCGCATCGAAACAGATGAGCACTTCAAAGCGGCGAAGTGAA	540						
DB	481	ACTTGTAACGTGTAATGGGCGCATCGAAACAGATGAGCACTTCAAAGCGGCGAAGTGAA	540						
QY	541	ATTTTGGTTCGATAAAGACATCGAGCTCAATGCAAGCAAGAGAGTAACCGAATTTAGGT	600						
DB	541	ATTTTGGTTCGATAAAGACATCGAGCTCAATGCAAGCAAGAGAGTAACCGAATTTAGGT	600						
QY	601	TACTAATGAAGGGCTTAAATCCTTGCAATGTGGGTAGCCATTTCCACTTTCTTGAAGCTAA	660						
DB	601	TACTAATGAAGGGCTTAAATCCTTGCAATGTGGGTAGCCATTTCCACTTTCTTGAAGCTAA	660						
QY	661	CAAGGCCTAAATTCGATCGTGAAGAAAGCCTATGGCAAGCCCTAGATATTCCTCTGG	720						
DB	661	CAAGGCCTAAATTCGATCGTGAAGAAAGCCTATGGCAAGCCCTAGATATTCCTCTGG	720						
QY	721	CAACACGCTACGATTCGGGCGAGGACAAACCCGCAAGTGCAGTTGATTCCTCTGGTGG	780						
DB	721	CAACACGCTACGATTCGGGCGAGGACAAACCCGCAAGTGCAGTTGATTCCTCTGGTGG	780						
QY	781	CAGTAAAAAAGTGAATTCGGCATGAACGGCTTGTGAATAACATTCGGGATGAACGCCATAA	840						
DB	781	CAGTAAAAAAGTGAATTCGGCATGAACGGCTTGTGAATAACATTCGGGATGAACGCCATAA	840						
QY	841	ACATAAAGCGCTTGCAAGGCGAAATCTCAGGATTTTATCAAGTAAGGAGACTCCCATGA	900						
DB	841	ACATAAAGCGCTTGCAAGGCGAAATCTCAGGATTTTATCAAGTAAGGAGACTCCCATGA	900						
QY	901	AAATGAAAAACAGATAATGTAATACCTACGACCCCAACCAAGGCGATAAAGTGCCT	960						
DB	901	AAATGAAAAACAGATAATGTAATACCTACGACCCCAACCAAGGCGATAAAGTGCCT	960						
QY	961	TAGGAGATACCGATCTTTGGGCGAAGTAGAAACATGACTATATACCACCTATGGCGAAGAAC	1020						
DB	961	TAGGAGATACCGATCTTTGGGCGAAGTAGAAACATGACTATATACCACCTATGGCGAAGAAC	1020						
QY	1021	TTAAATTTGGCGGGTAAAACTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGATG	1080						
DB	1021	TTAAATTTGGCGGGTAAAACTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGATG	1080						
QY	1081	AAAAACCCCTAGATTTAGTTCATCACTAACCGCATGATTAATCGACTACACCGGATTTACA	1140						
DB	1081	AAAAACCCCTAGATTTAGTTCATCACTAACCGCATGATTAATCGACTACACCGGATTTACA	1140						
QY	1141	AAGCCGACATTTGGATTTAAAAACGGCAAAATCCATGCAATGGCAAGCAAGAAACAGG	1200						
DB	1141	AAGCCGACATTTGGATTTAAAAACGGCAAAATCCATGCAATGGCAAGCAAGAAACAGG	1200						
QY	1201	ACATGCAAGATGGCGTAAAGCCCTCATATGGTCGTGGGTGTGGGCAAGCAAGCACTAGCAG	1260						
DB	1201	ACATGCAAGATGGCGTAAAGCCCTCATATGGTCGTGGGTGTGGGCAAGCAAGCACTAGCAG	1260						
QY	1261	GGGAAGGTATGATTTATACCGCTGGGGAAATCGATTCACACCCACCTTCCTTTCTCCAC	1320						
DB	1261	GGGAAGGTATGATTTATACCGCTGGGGAAATCGATTCACACCCACCTTCCTTTCTCCAC	1320						
QY	1321	AACAAATTCCTACCGCTTAGCCCAATGCGCTTACAAACCATGTTTGGAGCGGCAAGGTC	1380						
DB	1321	AACAAATTCCTACCGCTTAGCCCAATGCGCTTACAAACCATGTTTGGAGCGGCAAGGTC	1380						
QY	1381	CTGTAGATGGCAGCAATGCGACTACTACTCTCGGGCAAAATGGAATTTGCACCCCATGT	1440						
DB	1381	CTGTAGATGGCAGCAATGCGACTACTACTCTCGGGCAAAATGGAATTTGCACCCCATGT	1440						
QY	1441	TGCGCGCAGCAAGAGATTTCTATGATGTGGGCTTTTGGGCAAAAGCAATAGCTCTA	1500						
DB	1441	TGCGCGCAGCAAGAGATTTCTATGATGTGGGCTTTTGGGCAAAAGCAATAGCTCTA	1500						
QY	1501	GCAAAAAACAACTTTGTAGAACAGTAGAAGCGGGCGGATTTGGTTTAAATTTGCATGAAG	1560						
DB	1501	GCAAAAAACAACTTTGTAGAACAGTAGAAGCGGGCGGATTTGGTTTAAATTTGCATGAAG	1560						
QY	1561	ACTGGGCGACAAACCAAGTCGATCGATCTCTCTGAGCGTGGCAGATGAATAGCATG	1620						
DB	1561	ACTGGGCGACAAACCAAGTCGATCGATCTCTCTGAGCGTGGCAGATGAATAGCATG	1620						
QY	1621	TGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCAAGGTTATGTAGATGACACCTTAA	1680						
DB	1621	TGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCAAGGTTATGTAGATGACACCTTAA	1680						
QY	1681	ATGCAATGAACGGGCGGCCATCCATGCTACCAATTTAGGAGCGGGTGGAGGACACT	1740						
DB	1681	ATGCAATGAACGGGCGGCCATCCATGCTACCAATTTAGGAGCGGGTGGAGGACACT	1740						
QY	1741	CACCTGATTTTATCACCATGGCAGCGAGCTCAATATTTCTACCTCTCCACACCCCA	1800						
DB	1741	CACCTGATTTTATCACCATGGCAGCGAGCTCAATATTTCTACCTCTCCACACCCCA	1800						
QY	1801	CTATTCCCTATACCAATTAATACGGTTGAGAACACTTTAGACATGCTCATGATGCCACC	1860						
DB	1801	CTATTCCCTATACCAATTAATACGGTTGAGAACACTTTAGACATGCTCATGATGCCACC	1860						
QY	1861	ACCTAGACAAACCGATCCGAGGATTTTCAATTTTCTCAAAGCGGTATCCGCCCGGCT	1920						
DB	1861	ACCTAGACAAACCGATCCGAGGATTTTCAATTTTCTCAAAGCGGTATCCGCCCGGCT	1920						
QY	1921	CTATCGGGCTGAAGATGTCTCCATGATATGGGTGTGATCGCGATGACAAGCTCGGAT	1980						
DB	1921	CTATCGGGCTGAAGATGTCTCCATGATATGGGTGTGATCGCGATGACAAGCTCGGAT	1980						
QY	1981	CGCAAGCAATGGGCGTTCAGCGAAGTGAATTCCTCGAATTTGGCAGACTTCGGGATAAGA	2040						
DB	1981	CGCAAGCAATGGGCGTTCAGCGAAGTGAATTCCTCGAATTTGGCAGACTTCGGGATAAGA	2040						
QY	2041	ATAAAAAAGAAATTTGGTAAAGCTTCTGGAAGATGGCAAGATAACGATAATTTCCGATTA	2100						
DB	2041	ATAAAAAAGAAATTTGGTAAAGCTTCTGGAAGATGGCAAGATAACGATAATTTCCGATTA	2100						

QY 481 ACTTGTAACCTGTGAATTTGGCCCATCGAACCCAGATGAGCACTTCAAAAGCGGGCGAAGTGAA 540
DB 481 ACTTGTAACTGTGAATTTGGCCCATCGAACCCAGATGAGCACTTCAAAAGCGGGCGAAGTGAA 540
QY 541 ATTGTGTCGATAAAGACATCGAGCTCAATGCGAGCAAGAGTAAACCGAATCTGAGGT 600
DB 541 ATTGTGTCGATAAAGACATCGAGCTCAATGCGAGCAAGAGTAAACCGAATCTGAGGT 600
QY 601 TACTAATGAAGGGCCCTAAATCTTGTGATGTGGGTAGCCATTTCCACTCTTTTGAAGCTAA 660
DB 601 TACTAATGAAGGGCCCTAAATCTTGTGATGTGGGTAGCCATTTCCACTCTTTTGAAGCTAA 660
QY 661 CAAGGCATTAATAATTCGATCGTGAATAAGCGCTATGCGAAACGCGCTAGATATCCCTCTGG 720
DB 661 CAAGGCATTAATAATTCGATCGTGAATAAGCGCTATGCGAAACGCGCTAGATATCCCTCTGG 720
QY 721 CAACCGCTACGATTTGGGCGAGGACAAACCCCGCAAGTGCAGTTCGATTCCTCTTGGTGG 780
DB 721 CAACCGCTACGATTTGGGCGAGGACAAACCCCGCAAGTGCAGTTCGATTCCTCTTGGTGG 780
QY 781 CAGTAAAAAGTGAATTTGGCATGAACGGGCTTGTGAATAACATCGCGGATGAACGCCATAA 840
DB 781 CAGTAAAAAGTGAATTTGGCATGAACGGGCTTGTGAATAACATCGCGGATGAACGCCATAA 840
QY 841 ACATAAGCGCTTGACAAGCGGAAATCTCACCGATTTTCAAGTAAGGAGACTCCCATGA 900
DB 841 ACATAAGCGCTTGACAAGCGGAAATCTCACCGATTTTCAAGTAAGGAGACTCCCATGA 900
QY 901 AANATGAANAACAAAGATATGTAATACCTTACCGACCCCAAGGCGATATAAGTGGCT 960
DB 901 AANATGAANAACAAAGATATGTAATACCTTACCGACCCCAAGGCGATATAAGTGGCT 960
QY 961 TAGGAGATACCGATCTTTGGGCGAGAAGTAGAAATGACTATACCACTATGGCGGAAGAAC 1020
DB 961 TAGGAGATACCGATCTTTGGGCGAGAAGTAGAAATGACTATACCACTATGGCGGAAGAAC 1020
QY 1021 TTAATTTGGCGGGTAAATCTATCCGTGAGGATATGGGTACAGACAATAGCCCTGATG 1080
DB 1021 TTAATTTGGCGGGTAAATCTATCCGTGAGGATATGGGTACAGACAATAGCCCTGATG 1080
QY 1081 AAACACCCCTAGATTTAGTATCATCACTAAACCGGATGATTATCGACTACACCGGATTTTACA 1140
DB 1081 AAACACCCCTAGATTTAGTATCATCACTAAACCGGATGATTATCGACTACACCGGATTTTACA 1140
QY 1141 AAGCCGACATTTGGATTTAAAAACGGCAAAATCCATGGCATTTGGCAAGGCAAGGAAACAAGG 1200
DB 1141 AAGCCGACATTTGGATTTAAAAACGGCAAAATCCATGGCATTTGGCAAGGCAAGGAAACAAGG 1200
QY 1201 ACATGCAAGATGGCGTAAGCCCTCATATGGTTCGTGGGTGGGCGACAGAACACTAGCAG 1260
DB 1201 ACATGCAAGATGGCGTAAGCCCTCATATGGTTCGTGGGTGGGCGACAGAACACTAGCAG 1260
QY 1261 GGAAGGTATGATTTATACCGCTGGGGGAATCGATTCACATCACACCACTCTCTTTCTCCAC 1320
DB 1261 GGAAGGTATGATTTATACCGCTGGGGGAATCGATTCACATCACACCACTCTCTTTCTCCAC 1320
QY 1321 AACATTTCCCTACCGCTCTAGCCAAATGGCGGTTAACACCATGTTTGGAGGGCGGACAGGTC 1380
DB 1321 AACATTTCCCTACCGCTCTAGCCAAATGGCGGTTAACACCATGTTTGGAGGGCGGACAGGTC 1380
QY 1381 CTGTAGATGGCAGCAATGCGACTACTATCACTCCGGGCAAAATGGAACTTGACCGCATGT 1440
DB 1381 CTGTAGATGGCAGCAATGCGACTACTATCACTCCGGGCAAAATGGAACTTGACCGCATGT 1440
QY 1441 TGGCGCAGCAGAGAGATATTTATGAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTA 1500
DB 1441 TGGCGCAGCAGAGAGATATTTATGAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTA 1500
QY 1501 GCAAAAACAACTTTGTAGAACAGTAGAAGGGCGGATTTGGTTTAAATTTGATGAAG 1560
DB 1501 GCAAAAACAACTTTGTAGAACAGTAGAAGGGCGGATTTGGTTTAAATTTGATGAAG 1560
QY 1561 ACTGGGGCAACACCAAGTGGATCGATCACTGCTTGGCGGTGGCAGATGAATACGATG 1620

DB 1561 ACTGGGGCAACACCAAGTGGATCGATCACTGCTTGAAGCGTGGCAGATGAATACGATG 1620
QY 1621 TGCAAGTTTGTATTCACACCGATACAGTCAATGAGGAGGTTATGTAGATGACACCCCTAA 1680
DB 1621 TGCAAGTTTGTATTCACACCGATACAGTCAATGAGGAGGTTATGTAGATGACACCCCTAA 1680
QY 1681 ATGCAATGAACGGGCGGCGCATCGATGCTTACACATGAGGAGGCGGCTGAGAGACT 1740
DB 1681 ATGCAATGAACGGGCGGCGCATCGATGCTTACACATGAGGAGGCGGCTGAGAGACT 1740
QY 1741 CACTGATGTTATACCATTAATACCGTTGAGGAGCTCAATATTTCTACCCCTCTCCACACCCCA 1800
DB 1741 CACTGATGTTATACCATTAATACCGTTGAGGAGCTCAATATTTCTACCCCTCTCCACACCCCA 1800
QY 1801 CTATTTCCCTATACCATTAATACCGTTGAGGAGCTCAATATTTCTAGACATGCTCATGACATGCCACC 1860
DB 1801 CTATTTCCCTATACCATTAATACCGTTGAGGAGCTCAATATTTCTAGACATGCTCATGACATGCCACC 1860
QY 1861 ACCTAGACAAACCGCATCCGCGAGGATTTACAAATTTCTCAAGCGGTATCCGCCCGGCT 1920
DB 1861 ACCTAGACAAACCGCATCCGCGAGGATTTACAAATTTCTCAAGCGGTATCCGCCCGGCT 1920
QY 1921 CTATCGCGCTGAAGATGTCTCCATGATATGGGTGTGATCGCGATGACAAAGCTCGGAT 1980
DB 1921 CTATCGCGCTGAAGATGTCTCCATGATATGGGTGTGATCGCGATGACAAAGCTCGGAT 1980
QY 1981 CGCAAGCAATGGGCGGTGCGAGGAGTGTCTCGAACTTTGGCAGACTCGGATAAGA 2040
DB 1981 CGCAAGCAATGGGCGGTGCGAGGAGTGTCTCGAACTTTGGCAGACTCGGATAAGA 2040
QY 2041 ATAAAAAAGAAATTTGGTAACTTCTCGAAGATGCGAAAGATAACGATTAATTTCCGCAATTA 2100
DB 2041 ATAAAAAAGAAATTTGGTAACTTCTCGAAGATGCGAAAGATAACGATTAATTTCCGCAATTA 2100
QY 2101 AGCGCTACATCTCAAAATACATTAACCCCGCTTTGACCCACGGCGTGAGCGAGTATA 2160
DB 2101 AGCGCTACATCTCAAAATACATTAACCCCGCTTTGACCCACGGCGTGAGCGAGTATA 2160
QY 2161 TCGGCTCTGTGGAAGAGGCAAGATCGCGACTTTGGTGTGTGGAATCCTGCTTTTGTG 2220
DB 2161 TCGGCTCTGTGGAAGAGGCAAGATCGCGACTTTGGTGTGTGGAATCCTGCTTTTGTG 2220
QY 2221 GCGTAAAAACCAAAATCGTATCAAAAGCGGTATGTGTGCTTCTCTGAAATGGCGAT 2280
DB 2221 GCGTAAAAACCAAAATCGTATCAAAAGCGGTATGTGTGCTTCTCTGAAATGGCGAT 2280
QY 2281 CTAAACGGCTCTGTGCCCATCTCCCAACCGGTTTATTACCGGAAATGTTTGGGATCAG 2340
DB 2281 CTAAACGGCTCTGTGCCCATCTCCCAACCGGTTTATTACCGGAAATGTTTGGGATCAG 2340
QY 2341 GCAAGGCAAAATTTGACACAGCATCACTTTTGTTCCAAAGTCGCTATGAAATGGCG 2400
DB 2341 GCAAGGCAAAATTTGACACAGCATCACTTTTGTTCCAAAGTCGCTATGAAATGGCG 2400
QY 2401 TGAAAGAAAAGCTGGGCTTTAGAGCGCAAGTTCTACCGGTCAAAAATGCGGTAAACATCA 2460
DB 2401 TGAAAGAAAAGCTGGGCTTTAGAGCGCAAGTTCTACCGGTCAAAAATGCGGTAAACATCA 2460
QY 2461 CCAAGAAAGACTTCAAGTTCAACGACAAAACGGCAAAATCAACCGTCAATTCGAAACCT 2520
DB 2461 CCAAGAAAGACTTCAAGTTCAACGACAAAACGGCAAAATCAACCGTCAATTCGAAACCT 2520
QY 2521 TCGAGGCTTTGTAGATGGCAAACTCTGCACCTCTTAACCCACCTCGCAAGTGCCTTAG 2580
DB 2521 TCGAGGCTTTGTAGATGGCAAACTCTGCACCTCTTAACCCACCTCGCAAGTGCCTTAG 2580
QY 2581 CCCAGCGCTACACTTTTCTTAGGCACAATGCCCTTTGGGGCAGGTTATTTTAGGAA 2640
DB 2581 CCCAGCGCTACACTTTTCTTAGGCACAATGCCCTTTGGGGCAGGTTATTTTAGGAA 2640
QY 2641 TCTTTCATCAACGCAACCTGCAATCGGTCTTTCGTGTGCGATCGTGTGCTTTTAAACAAAC 2700

Db	2641	TCCTCATCAACGCCACCTGCAATCGGTCTTGCGTGTGCGGATCGTGCCTTTAAAAAAC	2700
Qy	2701	TTTTCACTTTTAAGCAATGCCATTTTAAATTAATTAATTTCTTATAATTAATTAATATAT	2760
Db	2701	TTTTCACTTTTAAGCAATGCCATTTTAAATTAATTAATTTCTTATAATTAATTAATATAT	2760
Qy	2761	TATGCCCTCATTTTTTAAAGGAGAAATATATGCTAGGCTTTTGGTATATGCTATGTGGGG	2820
Db	2761	TATGCCCTCATTTTTTAAAGGAGAAATATATGCTAGGCTTTTGGTATATGCTATGTGGGG	2820
Qy	2821	TTGTTTGGTCTGGGCGCAAGGGTATGAAACCCATCGCCTCAAAAAAGTAGAAGCCAC	2880
Db	2821	TTGTTTGGTCTGGGCGCAAGGGTATGAAACCCATCGCCTCAAAAAAGTAGAAGCCAC	2880
Qy	2881	AGG 2883	
Db	2881	AGG 2883	
RESULT 3			
BD185306			
LOCUS	BD185306	2452 bp	DNA linear PAT 17-JUN-2003
DEFINITION	Helicobacter felis vaccine.		
ACCESSION	BD185306		
VERSION	BD185306.1	GI:31877506	
KEYWORDS	JP 2002355054-A/5.		
SOURCE	Helicobacter felis		
ORGANISM	Helicobacter felis		
REFERENCE	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.		
AUTHORS	Kusters, J.G. and Cattoli, G.		
TITLE	Helicobacter felis vaccine		
JOURNAL	Patent: JP 2002355054-A 5 10-DEC-2002;		
COMMENT	AKZO NOBEL NV		
OS	Helicobacter felis		
PN	JP 2002355054-A/5		
PD	10-DEC-2002		
PP	16-JUL-2001 JP 2001214711		
PR	17-JUL-2000 EP 00202565.8		
PI	JOHANNES GERARDUS KUSTERS, GIOVANNI CATTOLI		
PC	C12N15/09, A61K38/00, A61K39/106, A61K39/118, A61K39/12, A61K39/175, PC A61K39/23,		
PC	A61K39/235, A61K39/39, A61K39/395, A61P1/04, A61P31/04, C12N1/15,		
PC	C12N1/19,		
PC	C12N1/21, C12N5/10, C12N9/80, C12Q1/68, G01N33/15, G01N33/50, G01N33/PC 53,		
PC	G01N33/53, G01N33/566, G01N33/569/(C12N9/80, C12P1:01), (C12Q1/68, PC C12P1:01),		
PC	C12N15/00, C12N5/00, A61K37/02		
CC	Helicobacter felis vaccine		
PH	Key	Location/Qualifiers	
FT	CDS	(49)..(728)	
FT	CDS	(739)..(2445).	
FEATURES	Location/Qualifiers		
source	1..2452		
	/organism="Helicobacter felis"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:214"		
ORIGIN			
Query Match	77.3%;	Score 2228;	DB 6; Length 2452;
Best Local Similarity	94.3%;	Pred. No. 0;	
Matches 2312;	Conservative 0;	Mismatches 140;	Indels 0; Gaps 0;
Qy	159	AGGACTATAATCCCATTTGCTTTAAATTTAAACAAAGGAGTAATAGTGAAACTCACAC	218
Db	1	AGGACTATAATCCCATTTGCTTTAAATTTAAACAAAGGAGTAATAGTGAAACTCACAC	60
Qy	219	CCAAAGACGAGAAAAGTTCTTGTTATATATGCGGGCGAAGTGGCTAGAAAGCGCAAG	278

Db	61	CCAAAGACGCAAGAAAAGTTCTTGTTATATATGCGGGCGAAGTGGCTAGAAAGCGCAAG	120
Qy	279	CAGAGGGCTTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGTGCCCATATTATGG	338
Db	121	CAGAGGGCTTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGTGCCCATATTATGG	180
Qy	339	ACGAAGCGCGCCGTGGAAAAAAAACCGTTGCCAGCTTATGGAAGAGTGCATGCACTTTT	398
Db	181	ACGAGCGCGTCTGTTGGCAAAAAACCGTTGCCAATTTATGGAAGAGTGTATGCACTTT	240
Qy	399	TGAAAAAGATGAAGTAATGCCCGGGTGGGTAATATGTTCCCGATCTAGGTGTAGAAG	458
Db	241	TGAAAAAGACGAGGTGATGCCCGGGTGGGGAATATGGTCCCTGATTTGGGCGTGGGAAG	300
Qy	459	CCACCTTTCTCTGATGGTACGAAACTTCTTAATCTGTAATTTGGCCCATCGAACAGATGAGC	518
Db	301	CCACTTTCCCGATGGCACCAAACTCTGTAATCTGTAATTTGGCCCATCGAACCTGATGAAC	360
Qy	519	ACTTCAAGCGGGCGAAGTGAAATTTGGTTGCCATAAAGACATCGAGCTCAATGCAGGCA	578
Db	361	ACTTTAAGGGGTGAAGTGAAATTTGGCTGTGATAAAGACATTGAACTCAACGCAGTA	420
Qy	579	AAGAAGTAACCGAACTTGAGGTTACTAATGAAGGCCCTAAATCCTTGCATGTGGGTAGCC	638
Db	421	AGGAAGTTACCGAACTAGAAGTTACTAAACGAAGGACCTAAATCCTTGCATGTGGGTAGCC	480
Qy	639	ATTTCCACTCTTTGAGCTAAACAGGCACCTAAATTCGATCGTGAAGAAAGCCCTATGGCA	698
Db	481	ATTTCCACTCTTTGAGCCAAACAAAGCATTGAAATTCGATCGGGAAGAAAGCCCTATGGCA	540
Qy	699	AACGCCCTAGATATTCCTCTGCGCAACACGCTAGCGATTGGGGCAGGACAAACCCGCAAG	758
Db	541	AAGCCCTAGATATTCCTCTGCGCAACACACTACGATTGGGGCAGGACAAACCCGTAAG	600
Qy	759	TGCAGTTGATTCTCTTGGTGGCAGTAAAAAGTGATTGGCATGAACGGGCTTGTGAATA	818
Db	601	TGCAGTTAATCTCTCTTGGCGGTAGTAAAAAGTGATTGGCATGAACGGGCTTGTGAATA	660
Qy	819	ACATCGCGATGAACGCCATAACATAAAGCGCTTGACAGGCGAATCTCACGGATTTA	878
Db	661	ATATTGCGGACGAAACGCCATAAACACAAAGCGCTAGACAAAGCAAAATCTCACGGATTTA	720
Qy	879	TCAAGTAAGGAGACTCCCATGAAAAATGAAAAACAAGATATGTAATACCTACGGACCC	938
Db	721	TCAAGTAAGGAGACTCCCATGAAAAATGAAAAACAAGATATGTAATACCTACGGACCC	780
Qy	939	ACCAAAGGCGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCGAAGTAGAACATGAC	998
Db	781	ACCACAGCGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCGAAGTAGAACATGAC	840
Qy	999	TATACACCTATGGCGAAGAACTTTAAATTTGGCGGGTAAAACTATCCGTGAGGGTATG	1058
Db	841	TATACACCTATGGCGAAGAACTTTAAATTTGGTGCAGTAAAACTATCCGTGAGGGTATG	900
Qy	1059	GGTCAGAGCAATAGCCCTGATGAAACACCCCTAGATTAGTTCATCTACGCGCATGATT	1118
Db	901	GGTCAGAGCAATAGCCCTGATGAAACACCCCTAGATTAGTTCATCTACGCGCATGATT	960
Qy	1119	ATCGACTACACCGGGATTTACAAAGCGCATTTGGGATTTAAAAACGGCAAAATCCATGGC	1178
Db	961	ATTGACTACACCGGGATTTACAAAGCGCATTTGGCATTAAAAATGCAAAATCCATGGC	1020
Qy	1179	ATTGCAAGGCGAGAAACAAAGCATGCAAGATGGCGTAAAGCCCTCATATGGTCTGGGT	1238
Db	1021	ATTGCAAGGCGAGAAACAAAGCATGCAAGATGGCGTAAAGCCCTCATATGGTCTGGGT	1080
Qy	1239	GTGGGCACAGAGCACTAGCAGGGGAGGTATGATTATTTACCGCTGGGGNAATCGATTCA	1298
Db	1081	GTGGGCACAGAGCACTAGCAGGGGAGGTATGATTATTTACCGCTGGGGNAATCGATTCA	1140
Qy	1299	CACCCCACTCTCTCTCCACAAATTCCTTACCGCTCTAGCCAAATGGCGTTTACAAAC	1358

Db 1141 CACACCCACTTCTCTCTCCACAAACAAATTCCCTTACCGCTCTAGCCAAATGCGGTTCACAA 1200
Qy 1359 ATGTTTGGAGCGGCACAGTCTCCTGTAGATGGCAGCAATGCGCACTACTATCACTCCGGGC 1418
Db 1201 ATGTTTGGCGGTGGCACAGCCCGGTAGATGGCAAGATGGCACTACCATCACTCCGGGC 1260
Qy 1419 AAATGGAACCTTGACCGCATGTTGCGGCGAGAGAGATATCTATGAATGTGGGCTTT 1478
Db 1261 AAATGGAACCTTGACCGCATGTTGCGGCGAGAGAGATATCTATGAATGTGGGCTTT 1320
Qy 1479 TTGGGCAAAAGGCAATAGCTCTAGCAAAAAACAATTGTGAACAAGTAGAAGCGGGCGG 1538
Db 1321 TTGGGCAAAAGGCAATAGCTCTAGTAAAAACAACCTTGTAGAACAGTAGAAGCGGGCGG 1380
Qy 1539 ATTGGTTTTAAATPGCATGAAGACTGGGGCACAAACCAAGTCGATCGATCACTGCTTG 1598
Db 1381 ATTGGTTTTAAATPGCATGAAGACTGGGGCACAACTCCAAGTCGATCGATCACTGCTTG 1440
Qy 1599 AGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACCCGATACAGTCAATGAGGCA 1658
Db 1441 AGCGTAGCAGATGAATACGATGTGCAAGTTTGTATACACCCGATACGCTCAATGAGGCA 1500
Qy 1659 GGTATGTAGATGACACCCCTAAATGCAATGAACGGCGGCCCATCCATGCTACCCACAT 1718
Db 1501 GGTATGTAGATGACACCCCTAAATGCAATGAACGGCGGCCCATCCATGCTACCCACAT 1560
Qy 1719 GAGGAGCGGTGGAGGACACTCACCTGATGTTTATCACCATGCGAGCGAGCTCAATATT 1778
Db 1561 GAGGAGCGGTGGAGGACACTCACCTGATGTTTATCACCATGCGAGCGAGTGAATATT 1620
Qy 1779 CTACCTCTCTCCACACCCCTCTATTCCTCTATACCAATTAATACGGTTGCGAGAACCTTA 1838
Db 1621 CTACCTCTCTCCACACCCCTCTATTCCTCTATACCAATTAATACGGTTGCGAGAACCTTA 1680
Qy 1839 GACATGCTCATGATGACACCCCTAGACAAAGCATCCGCGAGGATTTTACAATTTTCT 1998
Db 1681 GACATGCTCATGATGACACCCCTAGACAAAGCATCCGCGAGGATTTTCCAATTTTCT 1740
Qy 1899 CAAAGCGGTATCGCCCGCGCTCTATCGCGCTGAAGATGTGCTTCCATGATATGGGTGTG 1958
Db 1741 CAAAGCGGTATCGCCCGCGCTCTATCGCGCTGAAGATGTGCTTCCATGATATCGGTGTG 1800
Qy 1959 ATCCGATGACAAAGCTCGGATTCGCAAGCAATGGGCGTGCAGCGGAAGTATTCTCGA 2018
Db 1801 ATCCGATGACAAAGTTCGGATTTCGCAAGCAATGGGCGCTCGGGAAGTATTCTCTAGA 1860
Qy 2019 ACTTGGCAGACTGGGATAGAATAAATAAGAAATTTGGTAAGCTTCTCGAAGATGGCAAA 2078
Db 1861 ACTTGGCAAACTGCAGACAAGAAATAAATAAGAAATTTGGTAAGCTTCTCGAAGATGGTGA 1920
Qy 2079 GATAACGATTAATTTCCGCATTAAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTG 2138
Db 1921 GATAATGACAACTTCCGCATCAAAACGCTATATCTCCAAATACACCATTAATCCCGCTTTG 1980
Qy 2139 ACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGATCGCGACTTGGTG 2198
Db 1981 ACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGATCGCGACTTGGTG 2040
Qy 2199 GTGTGGAATCTGCTCTTTTGGCGTAAACCCCAAAATCTGTGATCAAGCGGTATGGTG 2258
Db 2041 GTGTGGAATCTGCTCTTTTGGCGTAAACCCCAAAATCTGTGATCAAGCGGTATGGTG 2100
Qy 2259 GTCTTCTCTGAAATGGGCGATTTCTAACCGCTCTGCGCCACTCCCGCAACCGGTTTATTAC 2318
Db 2101 GTGTCTCTGAAATGGGCGATTTCTAATGCGCTCTGTGCGCCACTCTCTAGCGCGTTATTATAC 2160
Qy 2319 CGCGAAATGTTTGGGCGATCAGCGCAAGCGCAAAATTTGACACAGCATCACTTTTGTCTCC 2378
Db 2161 CGCGAAATGTTTGGGCGATCAGCGCAAGCGCAAAATTTGACACAGCATCACTTTTGTCTCC 2220
Qy 2379 AAAGTCGCTATGAAATATGGCGTGAAGAAAGCTGGGCTTAGAGCGCCAAAGTCTACCG 2438
Db 2221 AAAGTCGCTATGAAATATGGTGTGAAGAAAGAACTAGGTTTATAGAGCGCAAGGTGCTCCCC 2280

Qy 2439 GTCAAAAACTGCCGTAAATCATCAACAGAAAGACTTCAAGTTCAACGACAAACCGCAAAA 2498
Db 2281 GTCAAAAACTGCCGTAAATCATCAACAGAGAGACTTCAAGTTCAACGACAAACCTGCAAAA 2340
Qy 2499 ATCACCGTGCATCCGAAAACCTTCGAGGTCTTTGTAGATGCGAAACTCTGCACCTCTAAA 2558
Db 2341 ATCACCGTGCATCCGAAAACCTTCGAGGTCTTTGTAGATGCGAAACTCTGCACCTCTAAA 2400
Qy 2559 CCCACCTCGCAAGTGCCTCTAGCCACGCTACACTTTCTTAGGCACAAT 2610
Db 2401 CCCACCTCGAAGTGCCTCTAGCCCAACGCTACACTTTCTTAGGCATAAT 2452

RESULT 4
AX356695
LOCUS AX356695 2452 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 13 from Patent EP1176192.
ACCESSION AX356695
VERSION AX356695.1 GI:18674032
KEYWORDS
SOURCE
ORGANISM
Helicobacter felis
Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.

REFERENCE 1
AUTHORS Kusters,J.G. and Cattoli,G.
TITLE Helicobacter felis vaccine
JOURNAL Patent: EP 1176192-A 13 30-JAN-2002;
Akzo Nobel N.V. (NL)

FEATURES
source
1. 2452
/organism="Helicobacter felis"
/mol_type="unassigned DNA"
/db_xref="taxon:214"
48..728
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23263.1"
/db_xref="GI:18674033"
/translation="MKLTPEQKFLIYAGEVARKKAEGLKLNQPEALAIYISAHIM
DEARRKKTVAELEMECHFLKDEVMGPMGVNMPDLEATFPDGFDDKLTIVNVPPIEP
DEHFKAQVFGCDKDIELNAGKEVTELVNEGFKSLHVGSHFFPFRANKALFKFRE
KAYGKRLDIPSGNTLRICAGQTRKVLIPGSGSKVIGMNLVNNIADERHKKALDK
AKSHGPTK"

CDS
739..2445
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23264.1"
/db_xref="GI:18674034"
/translation="MKMKQHYVNTYGTTCGDKVRLGDTDLWAEVEHDYTTVGBELKF
GAGTIREGMSQNSPDNTLDLVITNMIIDYGYKADIGIKNGKIHGIGKAGNKD
MQGVSPHMVVGTEALAGEMIIITAGGIDSHTHFLSPQFPPTALANGVITVMFGGT
GPVDTGNATITPGKMLHRAEYSMMVFLGKNSQSKQLVQVEAGATGFK
LHEDWGTTPSAIDHCLSVADYDVQVCIHITDVNEAGYVDVDTLANMNGRAIHAYHIEG
AGGHSPDVITMAGEVNTLPSSTTPTIPTYINTVAEHLDMLTCHHLDKRIREDLPFS
QSRIRPSIAEDVLHIGVIAMTSSDSQAGRAGEVIRPTMOTADKNKFKKLPEP
GADNDNPKRIKYITNPALTHGVSEYIGTVEEGKIADLVVWNPAPFGVKPKIVIK
LGMVVFSGMDSNASVPTPOPYVREMFHGHGAKFDTSITFVSKVAYENGVKELGL
ERKVLVKNCRNITKQDFKENDKAKITVDPKTFEYFVDGKLCSTSKPTSEVPLAQRYT
FF"

CDS
739..2445
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23264.1"
/db_xref="GI:18674034"
/translation="MKMKQHYVNTYGTTCGDKVRLGDTDLWAEVEHDYTTVGBELKF
GAGTIREGMSQNSPDNTLDLVITNMIIDYGYKADIGIKNGKIHGIGKAGNKD
MQGVSPHMVVGTEALAGEMIIITAGGIDSHTHFLSPQFPPTALANGVITVMFGGT
GPVDTGNATITPGKMLHRAEYSMMVFLGKNSQSKQLVQVEAGATGFK
LHEDWGTTPSAIDHCLSVADYDVQVCIHITDVNEAGYVDVDTLANMNGRAIHAYHIEG
AGGHSPDVITMAGEVNTLPSSTTPTIPTYINTVAEHLDMLTCHHLDKRIREDLPFS
QSRIRPSIAEDVLHIGVIAMTSSDSQAGRAGEVIRPTMOTADKNKFKKLPEP
GADNDNPKRIKYITNPALTHGVSEYIGTVEEGKIADLVVWNPAPFGVKPKIVIK
LGMVVFSGMDSNASVPTPOPYVREMFHGHGAKFDTSITFVSKVAYENGVKELGL
ERKVLVKNCRNITKQDFKENDKAKITVDPKTFEYFVDGKLCSTSKPTSEVPLAQRYT
FF"

ORIGIN
Query Match 77.3%; Score 2228; DB 6; Length 2452;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
Qy 159 AGGACTATAATCCCATTCGCTTAAATTTAAACAAGAGGTAATAGGTGAACACTCACAC 218
Db 1 AGGACTATAATCCCATTCGCTTAAATTTAAACAAGAGGTAATAGGTGAACACTCACAC 60

QY 219 CCAAAGAGCAAGAAAGTTCTTGTATATATATCGGGCGAAGTGGCTAGAAAGCGCAAG 278
DB 61 CCAAAGAGCAAGAAAGTTCTTGTATATATATCGGGCGAAGTGGCTAGAAAGCGCAAG 120
QY 279 CAGAGGGCTTAAAGCTCAACCAACCGAAGCCATTCCTTACATTAAGTGGCCCATATTATGG 338
DB 121 CAGAGGGCTTAAAGCTCAACCAACCGAAGCCATTCCTTACATTAAGTGGCCCATATTATGG 180
QY 339 ACGAAGCGGCGCGTGGAAAGAAACCGTTGCCAGCTTATGGAAGAGTGCATGCACATTTT 398
DB 181 ACGAGGCGGCTCGTGGCAAAAGAAACCGTTGCCGAACTTATGGAAGAGTGTATGCACATTTT 240
QY 399 TGAAGAAAGATGAAGTAATGCCCGGGGTGGGTAAATATGTTCCCGATCTAGAGTGTAGAAG 458
DB 241 TGAAGAAAGACAGAGTGATGCCCGGGGTGGGAAATATGGTCCCTGATTTGGGCGTGGAAAG 300
QY 459 CCACCTTTCTGATGTGACGAACCTTGAACCTGTGAATTTGGCCCATCGAACACGATGAGC 518
DB 301 CCACCTTTCCCGATGGCACCAAACTCGTAACCTGTGAATTTGGCCCATCGAACCTGATGAAC 360
QY 519 ACTTCAAAGCGGCGAAGTGAAATTTGGTTGCGATAAAGACATCGAGCTCAATGGCAGGCA 578
DB 361 ACTTTAAGCGGGTGAAGTGAAATTTGGCTGTGTATGAAGACATTTGAACCTCAAGCGAGGTA 420
QY 579 AAGAAGTAACCGAAGCTTGAGGTTTAAATGAAGGGCTTAAATCCTTTGCATGTGGGTAGCC 638
DB 421 AGGAAGTTACCGAAGCTAGAAAGTTACTAAAGGAGCACTTAAATCCTTTGCATGTGGGTAGCC 480
QY 639 ATTTCCACTTTCTTGAAGCTAAGGCACTAAGGCACTAAGGCTTGAATTCGATGAAAGGCTTATGGCA 698
DB 481 ATTTCCACTTTCTTGAAGCAACAAAGCACTTGAATTCGATCGGGAAGGCAATATGGCA 540
QY 699 AAGCGCTAGATATTCCTCTGGGCAACACGCTACGCAATGGGGCAGGACAAACCCGCAAG 758
DB 541 AAGCGCTAGATATTCCTCTGGGCAACACGCTACGCAATGGGGCAGGACAAACCCGTAAG 600
QY 759 TCGAGTTGATTCCTCTTGGTGGCAGTAAAGAAAGTGAATGGCATGAACGGGCTTGTGAATA 818
DB 601 TCGAGTTAATCCCTCTTGGCGGTAGTAAAGAAAGTGAATGGCATGAACGGGCTTGTGAATA 660
QY 819 ACATCGCGATGAACCGCATAAACATAAAGCGCTTGCAGAGGCGAAATCTCAGCGATTTA 878
DB 661 ATATTTGCGGCAAGCAAGCCATAACAAAGCGCTAGACAAAGCAAAATCTCAGCGATTTA 720
QY 879 TCAAGTAAAGGAGCTCCCATGAATAAAGAAACAAAGAAATATGTAATACCTTACGAGCCC 938
DB 721 TCAAGTAAAGGAGCTCCCATGAATAAAGAAACAAAGAAATATGTAATACCTTACGAGCCC 780
QY 939 ACCAAAGGCGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCAAGAGTAGAAACATGAC 998
DB 781 ACCAAGGCGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCAAGAGTAGAAACATGAC 840
QY 999 TATACACCTATGGGCAAGAACTTAAATTTGGCGCGGTAAACATACTATCGTGAAGGTATG 1058
DB 841 TATACACCTATGGGCAAGAACTTCAAAATTCGGTGCAGGTAAACATACTATCGTGAAGGTATG 900
QY 1059 GGTGAGCAATAGCCCTGATGAACACCCCTAGATTTAGTCATCACTAACCGGATGATT 1118
DB 901 GGTGAGCAATAGCCCTGATGAACACCCCTAGATTTAGTCATCACTAACCGGATGATT 960
QY 1119 ATCGACTACACCGGATTTTCAAAGCGGACATTTGGGATTTAAAGAAAGGCAAAATCCATGGC 1178
DB 961 ATTGACTACACCGGATTTTCAAAGCGGACATTTGGGATTTAAAGTGGCAAAATCCATGGC 1020
QY 1179 ATTGCAAGGCGAGGAAACAAAGGACATGCAAGATGGCGTAAAGCCCTCATATGTCGTGGGT 1238
DB 1021 ATTGCAAGGCGAGGAAACAAAGGACATGCAAGATGGCGTAAAGCCCTCATATGTCGTGGGT 1080
QY 1239 GTGGGCAAGAAGCACTACGAGGGAAGGTATGATTTATACCGCTGGGGAATCGATTCA 1298
DB 1081 GTGGGCAAGAAGCACTACGAGGGAAGGTATGATTTATACCGCTGGGGAATCGATTCA 1140
QY 1299 CACACCCACTTCTTCTTCCAAACAATTTCCCTACCGCTTAGCCCAATGGCGTTACAACC 1358

DB 1141 CACACCCACTTCTCTCTCCACAACAATTCCTTACCGCTCTAGCCAAATGGCGTTACAACA 1200
QY 1359 ATGTTTTGAGCGGCGCACAGAGTCTCTGTAGATGGCAGCAATGCGACTACTATCACTCCGGGC 1418
DB 1201 ATGTTTTGCGCGTGCCACAGGCCCGTAGATGGCAGCAATGCGACTACCATCACTCCGGGC 1260
QY 1419 AAATGGAACTTGCACCCGATGTTGCGGCGACGAGAAGATATTCTATGAATGTGGGCTTT 1478
DB 1261 AAATGGAACTTGCACCCGATGTTGCGGCGACGAGAAGATATTCTATGAATGTGGGCTTT 1320
QY 1479 TTGGGCAAGGCAATAGCTCTAGCAAAAGAAACAACTTGTAGAACAAAGTAGAAGCGGCGCG 1538
DB 1321 TTGGGCAAGGCAATAGCTCTAGTAAAAAACAACCTTGTAGAACAAAGTAGAAGCGGCGCG 1380
QY 1539 ATTGTTTTAAATTTGCAATGAAGACTGGGGCAACAACCAAGTGGCGATCGATCACTGCTTG 1598
DB 1381 ATTGTTTTAAATTTGCAATGAAGACTGGGGCAACAACCTTGTAGAACAAAGTAGAAGCGGCGCG 1440
QY 1599 AGCGTGGCAGATGAATACGATGTGCAAGTGTGTATCCACCCGATACAGTCAATGAGGCA 1658
DB 1441 AGCGTAGCAGATGAATACGATGTGCAAGTGTGTATACACCCGATACGCTCAATGAGGCA 1500
QY 1659 GGTATGTAGATGACACCCCTAAATGCAATGAACGGGCGGCGCATCCATGCCTACCAATTT 1718
DB 1501 GGTATGTAGATGACACCCCTAAATGCAATGAACGGGCGGCGCATCCATGCCTACCAATTT 1560
QY 1719 GAGGAGCGGTGGAGGACACTCACCTGATGTATACCATGTGGAGGCGAGTCAATATT 1778
DB 1561 GAGGAGCGGTGGAGGACACTCACCTGATGTATACCATGTGGAGGCGAAGTGAATATT 1620
QY 1779 CTACCTCTCTCCACACCCCTACTTATCCCTATACCATTAATACGGTTGCGAGAACACTTTA 1838
DB 1621 CTACCTCTCTCCACACCCCTACTTATCCCTATACCATTAATACGGTTGCGAGAACACTTTA 1680
QY 1839 GACATGCTATGACATGCCACCCCTAGACAAACGCAATCCGCGAGGATTTACAATTTCT 1898
DB 1681 GACATGCTTATGACCTGCCACCCCTAGATAAAGCATCCGAGGATCTCCAATTTCT 1740
QY 1899 CAAAGCGGTATCCGCCCGGCTCTATCGCGCTCAAGATGTCTCCATGATATGGGTGTG 1958
DB 1741 CAAAGCGGTATCCGCCCGGCTCTATCGCGCTCAAGATGTCTCCATGATATGGGTGTG 1800
QY 1959 ATCGCATGACAAAGCTCGGATTCGCAAGCAATGGGCGTGCAAGCAAGTGAATCTCTCGA 2018
DB 1801 ATCGCATGACAAAGTTCGATTCGCAAGCAATGGGCGCTGGGGAAGTGAATCTCTAGA 1860
QY 2019 ACTTGGCAGACTCGGATAAGAAATAAAAAAGAAATTTGGTAAGCTTCTTGAAGATGGCAAA 2078
DB 1861 ACTTGGCAAACTGCAGACAAAGAAATAAAAAAGAAATTTGGTAAGCTTCTTGAAGATGGTGA 1920
QY 2079 GATAACGATTAATTCGCGATTAGCGCTACATCTCCAAATACATATCAACCCCGCTTTG 2138
DB 1921 GATAATGACAACTTCCGCTCAAAACGCTATATCTCCAAATACACCAATTAATCCCGCTTTG 1980
QY 2139 ACCACGGCTGAGCGATATATCGGCTCTGTGGAAGAGGCAAGATCGCGACTTGGTG 2198
DB 1981 ACCATGGCTGAGCGATATATCGGCTCTGTGGAAGAGGCAAGATCGCGACTTGGTG 2040
QY 2199 GTGTGGAATCCTGCCCTTTTGGCGTAAAAACCCAAAAATTCGTGATCAAGGCGGTATGGTG 2258
DB 2041 GTGTGGAATCCTGCCCTTTTGGCGTAAAAACCCAAAAATTCGTGATCAAGGCGGTATGGTG 2100
QY 2259 GTCTTCTGAAATGGCGATTTCTAACCGCTGTGTGCCACTCCCCAAACCGGTTATTATAC 2318
DB 2101 GTCTTCTGAAATGGCGATTTCTTAATTCGCTGTGTGCCCACTCCTCGACCGGTTATTATAC 2160
QY 2319 CGGAAATGTTTTGGGATCACGCAAGCGAAATTTTGACACAGCATCACATTTTGTGTTC 2378
DB 2161 CGGAAATGTTTTGGGATCACGCAAGCGAAATTTTGACACAGCATCACATTTTGTGTTC 2220
QY 2379 AAAGTGGCTTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAGCGGCAAGTTCTTACCG 2438


```
Db 2221 AAAGTCGCCTATGAAATGGTGAAGAAAGAAACTAGGTTTAGAGCGCAAGGTGCTCCCC 2280
Qy 2439 GTCAAAAACTCGCGTAACATCACCAAGAAAGACTTCAAGTTCAACGACCAAAACGGCAAAA 2498
Db 2281 GTGAAAAACCTGCGTAACATCACCAAGAGGACTTCAAGTTCAACGACCAAAACCTGCAAAA 2340
Qy 2499 ATCACCGTCGATCGAAACCTTCGAGGCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2558
Db 2341 ATCACCGTCGATCGAAACCTTCGAGGCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2400
Qy 2559 CCACCTCGCAAGTGCCTCTAGCCCGAGCGCTACACTTCTTCTAGGCACAAT 2610
Db 2401 CCACCTCTCAAGTGCCTCTAGCCCAACGCTACACTTCTTCTAGGCATAAT 2452

RESULT 5
BD185303
LOCUS Helicobacter felis vaccine.
DEFINITION
ACCESSION BD185303
VERSION BD185303.1 GI:31877503
KEYWORDS JP 2002355054-A/2.
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 2405)
AUTHORS Kusters,J.G. and Cattoli,G.
TITLE Helicobacter felis vaccine
JOURNAL Patent: JP 2002355054-A 2 10-DEC-2002;
AKZO NOBEL NV
COMMENT OS Helicobacter felis
PN JP 2002355054-A/2
PD 10-DEC-2002
PF 16-JUL-2001 JP 2001214711
PR 17-JUL-2000 EP 00202565.8
PI JOHANNES GERARDUS KUSTERS, GIOVANNI CATTOLI
PC C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC
A61K39/23,
PC A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15,
PC C12N1/19,
PC C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/ PC
53,
PC G01N33/53,G01N33/566,G01N33/569/// (C12N9/80,C12R1/01), (C12Q1/68, PC
C12R1/01),
PC C12N15/00,C12N5/00,A61K37/02
CC Helicobacter felis vaccine
FH Key Location/Qualifiers
FT CDS (1)..(681)
FT CDS (692)..(2398).
FEATURES
source
1..2405
/organism="Helicobacter felis"
/mol_type="genomic DNA"
/db_xref="taxon:214"

ORIGIN
Query Match 74.9%; Score 2158.6; DB 6; Length 2405;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

Qy 206 GTGAACTCACCAAGAGCAAGAAAGTCTTGTATATATGCGGGCGAAGTGGCT 265
Db 1 GTGAAACTCACCAAGAGCAAGAAAGTCTTGTATATATGCGGGCGAAGTGGCT 60
Qy 266 AGAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGAAGCATTGCTTACATTAGT 325
Db 61 AGAAGCGCAAGCAGAGGGCTTAAAGCTCAATCAACCCGAAGCATTGCTTACATTAGT 120
Qy 326 GCCCATATTATGACGAAGCGCGCGTGGAAAAAACCCTTGCCAGCTTATGGAAGAG 385
```

```
Db 121 GCCCATATTATGACGAAGCGCGCGTGGAAAAAACCCTTGCTGAACTTATGGAAGAA 180
Qy 386 TGCATGCACATTTTGAATAAGATGAAGTAATCCCGGGGTGGTAAATATATGTTCCCGAT 445
Db 181 TGTATGCACATTTTGAATAAGATGAGTATGCCCGGTGTGGGAAATATGTTCCCTGAT 240
Qy 446 CTAGGTGTAGAACCACTTTCTGATGTGACGAAACTTGTAACTGTGAAATTTGGCCCATC 505
Db 241 TTGGCGGTAGAACCACTTTCCCGGTGGCAACCAACTCGTAACCGTGAATTTGGCCCAT 300
Qy 506 GAACCATGACGACACTTCAAAGCGGGGAGTGAATTTGGTTGCGATAAAGACATCGAG 565
Db 301 GAACCTGATGAACACTTTAAAGCGGTGAAGTGAATTTGGCTGTGATTAAGACATTTGAG 360
Qy 566 CTCAATGACGCAAGAAAGTAACCGAACTTGTAGGTTACTAATGAAGGCCCTAAATCTCTTG 625
Db 361 CTCAACGCGGTAAAGAGTTTACCGAGCTTGAAAGTTACCAACGAAGGACCTAAATCTCTTG 420
Qy 626 CATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAAACGAGCACTTAAATTCGATCGTGA 685
Db 421 CATGTGGGTAGCCATTTCCACTTCTTTGAAACCAACGAAGGCAATGAAATTCGATCGGAA 480
Qy 686 AAGCCCTATGCAAAACGCTAGATATTCCCTCTCGCAACGCTACGCAATTTGGGGCAGA 745
Db 481 AAGCCCTATGCAAAACGCTAGATATTCCCTCTCGCAACGCTACGCAATTTGGGGCAGA 540
Qy 746 CAAACCCGCAAAAGTGCAGTTGATTTCTCTTGGTGGCAGTAAATAAGTGAATTTGGCATGA 805
Db 541 CAAACCCGTAAGTGCAGTTAATCCCTCTTGGCGGTAGTAAATAAGTGAATTTGGCATGA 600
Qy 806 GGGCTTGTGAATAACATCGCGGATGAACGCAATAAACAATAAAGCGCTTGACAGGCGAAA 865
Db 601 GGGCTTGTGAATAATAATTGGCGAGCAACGCAATAAACAATAAAGCGCTTGACAGGCGAAA 660
Qy 866 TCTCACGGATTTATCAAGTAAAGGAGACTCCCATGAAATGAAATAAAGAGTATGTAAA 925
Db 661 TCTCACGGATTTATCAAGTAAAGGAGACTCCCATGAAATGAAATAAAGAGTATGTAAA 720
Qy 926 TACCTAGCGGACCCACCAAGCGATTAAGTGCCTTAGGAGATACCGATCTTTTGGGCAGA 985
Db 721 CACCTAGCGGACCCACCAAGCGATTAAGTGCCTTAGGAGATACCGATCTTTTGGGCAGA 780
Qy 986 AGTAGAATCATGACTATATACCATATGCGGAAGAACTTAAATTTGGCGGGGTAAACTAT 1045
Db 781 AGTAGAATCATGACTATATACCATATGCGGAAGAGCTCAATTTTGGCGGGGTAAACTAT 840
Qy 1046 CCGTGAGGTTATGGTTCAGAGCAATAGCCCTGATGAAACACCCCTAGATTTTGTAGTCATC 1105
Db 841 CCGTGAGGTTATGGTTCAGAGCAATAGCCCTGATGAAACACCCCTAGATTTTGTAGTCATC 900
Qy 1106 TAAACCGATGATTATCGACTACACCGGATTTTCAAAAGCCGACATTTGGGATTTAAACCG 1165
Db 901 CAACCGATGATTATCGACTACACCGGATTTTCAAAAGCCGACATTTGGTATTTAAATATG 960
Qy 1166 CAAATCCATCGCATTTGCCAAGGCAAGAAACAAGACATGCAAGATGCGGTAAAGCCCTCA 1225
Db 961 CAAATCCATCGCATTTGGCAAGGCGGGAACAAGACATGCAAGATGCGGTAAAGCCCTCA 1020
Qy 1226 TATGTCGTGGGTGTGGGCACAGAAGCACTAGCAGGGGAAGTATGATTTATACCGCTGG 1285
Db 1021 TATGTCGTGGGTGTGGGCACAGAAGCACTAGCAGGGGAAGTATGATTTATACCGCTGG 1080
Qy 1286 GGGAAATCGATTTCACACCCCACTTCTTTCTCCAAACAATTTCCCTACCGCTCTAGCCAA 1345
Db 1081 GGGGATCGATTTCGCACACCCCACTTCTCTCTCCCAACAATTTCCCTACCGCTCTAGCCAA 1140
Qy 1346 TGGCGTTACAAACCAATTTTGGAGGCGGCACAGGTCTCTGTAGATGGCAAGATGCGGACTAC 1405
Db 1141 TGGTGTACAAACCAATTTTGGAGGTGGCACAGGTCCGGGTAGATGGCAGCAAGATGCGGACCAC 1200
Qy 1406 TATCATCTCCGGGCAATGGAACCTTGACCGGATGTTGGCGGACGACGAGATTTCTAT 1465
```

Db 1201 CATCACTCGGGCAAAATGGAACCTTGACCGCATGTTGGCGCGAGCTGAAGAGTATTTCTAT 1260
Qy 1466 GAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAACAATTTGTAGAAACAAGT 1525
Db 1261 GAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCCAGTAAAAACAACACTCGTAGAACAGT 1320
Qy 1526 AGAAGGGGGCGGATGTTGGTTTAAATTTGAATGAAGACTGGGGCAAAACCAAGTGGCAT 1585
Db 1321 AGAAGGGGGCGGATGTTGGTTTAAATTTGAATGAAGACTGGGGCAAAACCAAGTGGCAT 1380
Qy 1586 CGATCACTCTTGAGCGTGCGAGATGAATACCATGTGCAAGTTTGTATCCACACCGATAC 1645
Db 1381 CGATCACTCTTGAGCGTAGAGATGAATACCATGTGCAAGTTTGTATCCACACCGATAC 1440
Qy 1646 AGTCAATGAGGAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGGCGGCCATCA 1705
Db 1441 GGTCAATGAGGAGGTTATGTAGATGACACCCCTAAATGCGATGAACGGGCGGCCATCA 1500
Qy 1706 TGCCTACCACTGAGGAGGCGGGTGGAGGACATCACTGATGTTATCACCATGGCAGG 1765
Db 1501 TGCCTACCACTGAGGAGGCGGGGAGGAGACACTCACCTGATGTTATCACCATGGCAGG 1560
Qy 1766 CGAGCTCAATATCTACCTCTCCACACCCCGCACTATTCCTATACCATTAATACGGT 1825
Db 1561 CGAGCTCAATATCTACCTCTCCACACCCCGCACTATTCCTATACCATTAATACGGT 1620
Qy 1826 TGCAGAACCTTAGACATGCTCATGACATGCGACACCTAGACAAACGATCCGCGAGG 1885
Db 1621 TGCAGAACCTTAGACATGCTCATGACATGCGACACCTAGACAAACGATCCGCGAGG 1680
Qy 1886 TTTACAATTTTCTCAAAGCGGTATCGGCCCGGCTCTATCGGGCTGGAAGATGTGTCCA 1945
Db 1681 TCTCCAGTTTTCCTCAAAGCGGTATCGGCCCGGCTCTATTCGGCTGGAAGATGTGTCCA 1740
Qy 1946 TGATATGGGTGATCGCGATGACAGCTCGGATTCGCAAGCAATGGGCGGTGCGAGGCA 2005
Db 1741 TGATATGGGTGATCGCGATGACAGCTCGGATTCGCAAGCAATGGGCGGTGCGAGGCA 1800
Qy 2006 AGTGATTCCTGCAACTTGGCAGACTCGGATGAAGATTAAGAAATTTGGTAACTTCC 2065
Db 1801 AGTGATTCCTGCAACTTGGCAGACTCGCAGCAAGAAATAAAGAAATTTGGTAACTTCC 1860
Qy 2066 TGAAGATGCCAAAGATACGATAAATTTCCGCAATTAAAGCGCTACATCTCCAAATACACTAT 2125
Db 1861 TGAAGATGGTCAGATATGACAACTTCGCGATCAACCGCTATATCTCCAAATACACTAT 1920
Qy 2126 CAACCCCGCTTTGACCCAGCGGTGAGCGATATATCGGCTCTGTGGAAGAGGGCAAGAT 2185
Db 1921 TAATCCCGCTTTGACCCATGGCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGAT 1980
Qy 2186 CGCGCACTTGGTGTGGAAATCCTGCTTTTGGCGTAAACCCCAAAATCGTGATCAA 2245
Db 1981 CGCGCACTTGGTGTGGAAATCCTGCTTTTGGTGTAAACCCCAAAATCGTGATCAA 2040
Qy 2246 AGCGGTATGGTGGTCTCTCTGAATGGGCGATTTCTAACCGGTCTGTGCCCACTCCCA 2305
Db 2041 AGCGGTATGGTGGTCTCTCTGAATGGGCGATTTCTAACCGGTCTGTGCCCACTCA 2100
Qy 2306 ACCGGTTTATACCGCAAAATGTTTGGGCATCACCGCAAGGGCGAAATTTGACACAGCAT 2365
Db 2101 GCGGTTTTATTACCGCAAAATGTTTGGGCATCACCGCAAGGGCGAAATTTGACACAGCAT 2160
Qy 2366 CACTTTTGTTCCTCAAGTGGCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAGG 2425
Db 2161 CACTTTTGTTCCTCAAGTGGCTATGAAATGGCGTGAAGAAAGAACTAGGCTTAGAGG 2220
Qy 2426 CCAAGTCTACCGGTCAAAATCTGCGGTAAACATCACCAAGAAAGCTTCAAGTCAACGA 2485
Db 2221 CAAGTGTCTACCGGTGAAGAACTGCGGCAACATCACTAGAAGAACTTCAAAATCAACAA 2280
Qy 2486 CAAAACGGCAAAATACACGTCGATCCGAAAACTTTCGAGGTCTTTGTAGATGGCAACT 2545
Db 2281 CAAGAGGGGCATATCACTGTGATCCTTAAACCTTCGAGGTCTTTGTAGATGGCAACT 2340

Qy 2546 CTGCACCTCTAAACCCACCTCGCAAGTGGCTCTAGGCCACCGCTACACTTCTTCTAGGC 2605
Db 2341 CTGCACCTCTAAACCCGCCTCTGAAGTGGCTCTAGGCCAACGCTACACTTCTTCTAGGC 2400
Qy 2606 ACAAT 2610
Db 2401 ACAAT 2405
RESULT 6
LOCUS AX356686 2405 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 4 from Patent EP1176192.
ACCESSION AX356686
VERSION AX356686.1 GI:18674023
KEYWORDS Helicobacter felis
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Helicobacteraceae; Helicobacter.
REFERENCE 1
AUTHORS Kusters,J.G. and Cattoli,G.
TITLE Helicobacter felis vaccine
JOURNAL Patent: EP 1176192-A 4 30-JAN-2002;
Akzo Nobel N.V. (NL)
FEATURES
Location/Qualifiers
1..2405
/organism="Helicobacter felis"
/mol_type="unassigned DNA"
/db_xref="taxon:214"
1..681
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23257.1"
/db_xref="GI:18674024"
/translation="MKLTPEQEKFLLYAGEVARKRKABGLKLNQPEAIAYISAHIM
DEARQKKTVAELMBECHMFLKDEVMGPGVMVVDLGVBEATPDGTLVTVMNPIEP
DHFVAGVEYKFGCKDIELNAGKEVTELVNKGPKSLHVGSHFHFETNKALKFDFRE
KAYKRLDIPSGNTLRIGAGTRKQVLIPLGGSKKVIQMGVLVNNIADERHKHKKALDK
AKSHGFIK"
692..2398
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23258.1"
/db_xref="GI:18674025"
/translation="MKMKQEVVNTYGPTTGDVKRLGDTDLMAEVEHDYTTYGBELKF
GAGKTIREGMGQSNPDENTLDLVITNMIIDYGIYKADIGIKNGKIHGIGKAGND
MODGVSPIHVVGVGTALAGEGMIITAGGIDSHTHFLSPQFPKALANGVTTFMGGCT
GVPDGNATTITPGKNLHRLRAAEYSNMVGFLOGNSSSKQLVEQVEAGAIQPK
LHEDNGTTPSAIDHCLSVADYDVQCIHTDVNEAGYVDDTLNMGRAIHAHYHISG
AGGHSIPDVTWAGELNLPSSSTPTPTINTVAEHLDMLTCHLDKRIKIRDLQFS
QSRIRPGSLAAEDVLHDIQVIAVTSQDSQAMGAGEVIPTWTQADKNKKEFGKLPED
GADNFRITKYSKITNPALTHGVSEYIGSVGEKIALDVVMNPAFGVQPKIVIK
GCMVFSMGSDNSASVPTPQPVYREMFHGHGKAKFDTSITFVSKVAYENGVEKELGI
ERKVLPVKNCRNITTKDFKFNKTAHITVDPFTFEVVDGKLCTSPASEVPLAQRIT
FF"

ORIGIN

Query Match 74.9%; Score 2158.6; DB 6; Length 2405;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
Qy 206 GTGAAACTCACACCCAAAGAGCAAGAAAAAGTTCTTGTATATTATGCGGGCGAAGTGGCT 265
Db 1 GTGAAACTCACACCCAAAGAGCAAGAAAAAGTTCTTGTATATTATGCGGGCGAAGTGGCT 60
Qy 266 AGAAGCGCAAGAGCAGAGGGCTTAAAGCTCAACCAACCGCAAGCCATTGCTTACATTAGT 325
Db 61 AGAAGCGCAAGAGCAGAGGGCTTAAAGCTCAATCAACCCGAGCCATTGCTTACATTAGT 120

QY 326 GCCATATATGAGCAAGCGCGCTGGAAAAAACCCTTCCCGAGCTTATGGAAGAG 385
Db 121 GCCATATATGAGCAAGCGCGCTGGAAAAAACCCTTCCCGAGCTTATGGAAGAG 180
QY 386 TGCATGCACCTTTTGAAGAAAGATGAAGTAATCCCGGGTGGTAAATGTTGCCGAT 445
Db 181 TGTATGCACCTTTTGAAGAAAGATGAAGTATGCCCGGTGGGGAATATGGTCCCTGAT 240
QY 446 CTAGTGTGAGAACCACTTTCCCTGATGTGACGAAACTTGTAACTGTGAATTTGGCCCATC 505
Db 241 TTGGCGGTAGAACCACTTTCCCGGATGGCACCAAACTCGTAACCGTGAATTTGGCCCAT 300
QY 506 GAAACAGATGAGCACTTCAAGCGGGGGAAGTGAATTTGGTGGCAGTAAGACATCGAG 565
Db 301 GAACTGTATGAAACACTTTAAAGCGGTGAAGTGAATTTGGCTGTGTATAAGACATTTGAG 360
QY 566 CTCAAATGCGAGCAAGTAACCGAACTTGAAGTGTACTAATCAAGGGCCCTAAATCCTTG 625
Db 361 CTCAACCGGGTAAGGAAGTTACCGAGCTTGAAGTTACCAACGAAGAACCTAAATCCTTG 420
QY 626 CATGTGGGTAGCCATTTCCACTTTCTTTGAAGCTTAACAAGGCACCTAAAATTCGATCGTGA 685
Db 421 CATGTGGGTAGCCATTTCCACTTTCTTTGAAGCAACAAGGCAATTTGAATTCGATCGGAA 480
QY 686 AAGCCTATGGGAAAACGCTAGATATATCCCTCTGGGCAACAGCTACGATTTGGGGCAGGA 745
Db 481 AAGCCTATGGGAAAACGCTAGATATATCCCTCTGGCAACAGCTACGATTTGGGGCAGGA 540
QY 746 CAACCCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAGATGATTGGCATGCAAC 805
Db 541 CAACCCGCAAGTGCAGTTGATTCCTCTTGGCGGTAGTAAAGATGATTGGCATGCAAC 600
QY 806 GGCCTTGTGAATAACATTCGCGGATGAACGCGCATAAACATAAAGCGCTTGACAAGGGGAAA 865
Db 601 GGCCTTGTGAATAATTTGCGGACGACGCGCATAAACATAAAGCACTAGACAGGCAAAA 660
QY 866 TCTCAGGATTTATCAAGTAAGGAGACTCCCATGAAATGAAACAAAGAAATATGTAAA 925
Db 661 TCTCAGGATTTATCAAGTAAGGAGACTCCCATGAAATGAAACAAAGAGTATGTAAA 720
QY 926 TACCTACGGACCCCAAGGCGATAAAGTGCCTTAGGAGATACCGATCTTTGGGCGA 985
Db 721 CACCTACGGACCCCAAGGCGATAAAGTGCCTTAGGAGATACCGATCTTTGGGCGA 780
QY 986 AGTAGAACATGACTATACCACTATGCGGAAGAACTTAAATTTGGCGGGGTAAAACTAT 1045
Db 781 AGTAGAACATGACTATACCACTATGCGGAAGAGCTCAAATTTGGCGGGGTAAAACTAT 840
QY 1046 CCGTAGGGGTATGGGTGAGAGCAATAGCCCTGTAGTAAACACCCCTAGATTTAGTCATC 1105
Db 841 CCGTAGGGGTATGGGTGAGAGCAATAGCCCGAGATGAAACACACCTTAGATTTAGTCATC 900
QY 1106 TAACCGGATGATTTCACTACACCGGGATTTCAAGCCGACATTTGGGATTTAAACGG 1165
Db 901 CAACCGGATGATTTCACTACACCGGGATTTTAAAGCCGACATTTGGTATTAAAAATGG 960
QY 1166 CAAAAATCCATGGCAATTCGCAAGGAGGAGAAACAAGGACATGCAAGTGCCTAGACCCCTCA 1225
Db 961 CAAAAATCCATGGTATTCGCAAGGCGGGGAAACAAAGACATGCAAGTGCCTAGACCCCTCA 1020
QY 1226 TATGGTCTGGGTGAGGCAAGAGCACTAGCAGGGGGAAGGTATGATTTATCCCGCTGG 1285
Db 1021 TATGGTCTGGGTGAGGCAAGAGCACTAGCAGGGGGAAGGTATGATTTATCCCGCTGG 1080
QY 1286 GGGATCGATTTACACACCACTTCTTTCTCAACAATTTCCCTACCGCTCTAGGCCAA 1345
Db 1081 GGGATCGATTTGACACACCACTTCTTTCTTTCCCAAAATTTCCCTACCGCTCTAGGCCAA 1140
QY 1346 TGGGTTTACACCACTGTTTGGAGCGGCACAGGTCCTGTAGATGCGCAAGTGCAGCTAC 1405
Db 1141 TGGGTTTACACCACTGTTTGGAGTGGCAAGGTCGCGGTAGATGCGCAAGTGCAGCTAC 1200
QY 1406 TATCACTCCGGGCAAAATGGAATTTGCACCGCATGTTGCGCGCAGCAGAGATTTCTAT 1465

Db 1201 CATCACTCCGGGCAAAATGGAACCTTGCAACCGCATTTGCGCGCAGCTGAAAGATTTCTAT 1260
QY 1466 GAATGTGGGCTTTTGGGCAAAAGCAATAGCTCTAGCAAAAAAACAATTTGTAGAACAGT 1525
Db 1261 GAATGTGGGCTTTTGGGCAAAAGCAATAGCTCCAGTAAAAAACAATCTGTAGAACAGT 1320
QY 1526 AGAAGCGGGCGCATTTGGTTTTAAATTTGCATGAAAGCTGGGGCAACAACAAGTGGAT 1585
Db 1321 AGAAGCGGGCGCATTTGGTTTTAAATTTGCATGAAAGCTGGGGCAACAACAAGTGGAT 1380
QY 1586 CGATCACTGCTTGCAGCGTGGCAGTGAATACGATGTGCAAGTTTGTATCCACCGGATAC 1645
Db 1381 CGATCACTGCTTGCAGCGTGGCAGTGAATACGATGTGCAAGTTTGTATCCACCGGATAC 1440
QY 1646 AGTCAATAGCGCAGGTTTATGTAGATGACACCTTAAATGCAATGAAACGGCGCGCATCCA 1705
Db 1441 GGTCAATAGCGCAGGTTTATGTAGATGACACCTTAAATGCGATGAAACGGCGCGCATCCA 1500
QY 1706 TGCCTACCACTTGAAGGAGCGGTGGAGACACTCACTGATGTTTATCAACATGGCAGG 1765
Db 1501 TGCCTACCACTTGAAGGAGCGGTGGAGACACTCACTGATGTTTATCAACATGGCAGG 1560
QY 1766 CGAGCTCAATATTTCTAACCCCTCTCCACCCACCTATTTCCCTATACCATTAATACGGT 1825
Db 1561 CGAGCTCAATATTTCTAACCCCTCTCCACCCACCTATTTCCCTATACCATTAATACGGT 1620
QY 1826 TGCAGAACACTTGAAGACATGCTCATGATGACACCACTAGACAAAGCATCCCGCAGGA 1885
Db 1621 TGCAGAACACTTGAAGACATGCTCATGATGACACCACTAGACAAAGCATCCCGCAGGA 1680
QY 1886 TTTTCAATTTTCTCAAAAGCGGTATTCGCCCGCGCTCTATCGCGCTGAAGATGTGCTCA 1945
Db 1681 TCTCCAGTTTTTCCAAAGCGGTATTCGCCCGCGCTCTATTCGCCGTGAAGATGTGCTCA 1740
QY 1946 TGATATGGGTGTGATTCGGATGACAAAGCTCGGATTCGCAAGCAATGGCGGTGAGGCGA 2005
Db 1741 TGATATTTGGGTGTGATTCGGATGACAAAGCTCGGATTCGCAAGCAATGGCGGTGAGGCGA 1800
QY 2006 AGTGATTTCTCGAACTTTGGCAGACTGCGGATGAAATAAAAAAGAAATTTGTAAGCTTCC 2065
Db 1801 AGTGATTTCTAGAACTTTGGCAACTGCAAGCAAGATAAAAAAGAAATTTGTAAGCTTCC 1860
QY 2066 TGAAGATGCAAAAGATAAACGATATTTCCGCAATTAAGCGCTACATCTCAAAATACACTAT 2125
Db 1861 TGAAGATGCTGACAGATTAATGACAACTTCGCACTCAACCGCTATATCTCCTAAATACACCAT 1920
QY 2126 CAACCCCGCTTTGACCCACCGCGTGAAGAGTATATCGGCTCTGTGGAAGGCGCAAGAT 2185
Db 1921 TAATCCCGCTTTGACCCATGCGAGTATATCGGCTCTGTGGAAGGCGCAAGAT 1980
QY 2186 CGCCGACTTGTGTGTGGAAATCCTGCTTTTGGCGTAAAAACCAAAATCGTGATCAA 2245
Db 1981 CGCCGACTTGTGTGTGGAAATCCTGCTTTTGGTGTAAAAACCAAAATCGTGATCAA 2040
QY 2246 AGCGGTATGGTGTCTTCTCTGAAATGGGCGATTTTAAACCGCTCTGTGCCACTCCCCA 2305
Db 2041 AGCGGTATGGTGTGTCTCTGAAATGGGCGATTTTAAACCGCTCTGTGCCACTCCCCA 2100
QY 2306 ACCGGTTTATTAACCGGAAATGTTTGGGCACTACGGCAAGCGGAAATTTGACACCAAGAT 2365
Db 2101 GCGCGTTTATTAACCGGAAATGTTTGGGCACTACGGCAAGCGGAAATTTGACACCAAGAT 2160
QY 2366 CACTTTTGTTCCTAAAGTCGCTTATGAAATGGGCTGAAAGAAAGCTGGGCTTAGAGCG 2425
Db 2161 CACTTTTGTTCCTAAAGTCGCTTATGAAATGGGCTGAAAGAAAGCTAGGCTTAGAGCG 2220
QY 2426 CCAAGTTTCTACCGGTCAAAAATCTCCGCTTAACATCAACAAAGAAAGCTTTCAAGTTCAACGA 2485
Db 2221 CAAGGTGCTACCGGTGAAAGAAATCTCCGCAACATCACTAAGAAAGAGCTTTCAAATTCACAA 2280
QY 2486 CAAAACGGCAAAATCAACCGGTGATTCGAAACCTTTCGAGTCTTTGTAGATGGCAAACT 2545

Db	2281	CARAGCGGCGCATATCACTCTCGATCTCTAAAACCTTCGAGGTCTTTGTAGTAGCGCAAAC	2340
Qy	2546	CTGCACCTCTAAACCACCCACTCGCAAGTGCCTCTAGCCAGCGCTACACTTTCTCTAGGC	2605
Db	2341	CTGCACCTCTAACCAGCGCCTCTGAAGTGCTCTAGGCCAACGCTACACTTTCTCTAGGC	2400
Qy	2606	ACAAT 2610 	
Db	2401	ACAAT 2405	
RESULT 7			
BD185305			
LOCUS			
DEFINITION		Helicobacter felis vaccine.	PAT 17-JUN-2003
ACCESSION		BD185305	
VERSION		BD185305.1 GI:31877505	
KEYWORDS		JP 2002355054-A/4.	
SOURCE		Helicobacter felis	
ORGANISM		Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales; Helicobacteraceae; Helicobacter. 1 (bases 1 to 2407)	
REFERENCE		Kusters,J.G. and Cattoli,G. Helicobacter felis vaccine Patent: JP 2002355054-A 4 10-DEC-2002;	
AUTHORS		AKZO NOBEL NV	
TITLE		OS Helicobacter felis	
JOURNAL		PN JP 2002355054-A/4	
COMMENT		PD JP 2002355054-A/4 PF 16-JUL-2001 JP 2001214711 PR 17-JUL-2000 EP 00202565.8 PI JOHANNES GERARDUS KUSTERS,GIOVANNI CATTOLI PC C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC A61K39/23, PC A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15, PC C12N1/19, PC C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/ PC 53, PC G01N33/53,G01N33/566,G01N33/569// (C12N9/80,C12R1:01) , (C12Q1/68, PC C12R1:01), PC C12N15/00,C12N5/00,A61K37/02 CC Helicobacter felis vaccine FH Key Location/Qualifiers FT CDS (2)..(682) FT CDS (693)..(2399).	
FEATURES		Location/Qualifiers	
source		1..2407 /organism="Helicobacter felis" /mol_type="genomic DNA" /db_xref="taxon:214"	
ORIGIN			
Query Match		74.9%; Score 2158.6; DB 6; Length 2407;	
Best Local Similarity		93.8%; Pred. No. 0;	
Matches 2251; Conservative		0; Mismatches 155; Indels 0; Gaps 0;	
Qy	206	GTGAACTCACACCCAAAGCAGGAGGTTAAAGCTCAACCAACCCGAGCCATTGCTTACATTAGT	265
Db	2	GTGAACTCACACCCAAAGCAGGAGGTTAAAGCTCAACCAACCCGAGCCATTGCTTACATTAGT	61
Qy	266	AGAAAGCGCAACGAGGGCTTAAAGCTCAACCAACCCGAGCCATTGCTTACATTAGT	325
Db	62	AGAAAGCGCAACGAGGGCTTAAAGCTCAACCAACCCGAGCCATTGCTTACATTAGT	121
Qy	326	GCCCATATTATGCAACGGCGCGTGGAAAAAACAACCGTTGCCAGCTTATGGAGAG	385
Db	122	GCCCATATTATGCAACGGCGCGCTGGAAAAAACAACCGTTGCCAGCTTATGGAGAG	181
Qy	386	TGCATGCACCTTTTGAATAAAGAATGAAGTAATATGCCCGGGTGGGTAAATATGTTCCCGAT	445

Db 1262 GAATGTGGGCTTTTGGGCAAGGCAATAGCTCCAGTAAGAAAAACAACATCTGTAGAACAAAT 1321
QY 1526 AGAAGCGGCGGCGATTTGGTTTAAATTTGCAATGAAGACTGGGGCACAAACCAAGTGCAT 1585
Db 1322 AGAAGCGGCGGCGATCGGCTTTAAATTTGCATGAAGACTGGGGCACAACTCCAAGTGAAT 1381
QY 1586 CGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC 1645
Db 1382 CGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC 1441
QY 1646 AGTCAATAGGCGAGGTATGTAGATGACACACCTAAATGCAATGAACGGGGCGGCCATCCA 1705
Db 1442 GGTCAATAGGCGAGGTATGTAGATGACACACCTAAATGCAATGAACGGGGCGGCCATCCA 1501
QY 1706 TGCTTACCACATTGAGGAGCGGGTGGAGGACACTCACCTGATGTTATCACCATGGCAGG 1765
Db 1502 TGCTTACCACATTGAGGAGCGGGGAGGACACTCACCTGATGTTATCACCATGGCAGG 1561
QY 1766 CGAGCTCAATATTCTACCTCTCCACACCCGCACTATTCCTTATACCATTAATACGGT 1825
Db 1562 CGAGCTCAATATTCTACCTCTCCACACCCGCACTATTCCTTATACCATTAATACGGT 1621
QY 1826 TGCAAGACCTTAGACATGCTATGATGATCGACATCGACCTAGACAAACGCAATCCGCGAGGA 1885
Db 1622 TGCAAGACCTTAGACATGCTATGATGATCGACCTAGACAAACGCAATCCGCGAGGA 1681
QY 1886 TTTTACAAATTTCTCAAAGCGGTATCGCCCGGCTCTATCGCGCTGAAGATGCTCTCA 1945
Db 1682 TTTTACAAATTTCTCAAAGCGGTATCGCCCGGCTCTATCGCGCTGAAGATGCTCTCA 1741
QY 1946 TGATATGGGTGATCGCGATGACAAGCTCGATTCGCAAGCAATGGGGCGTGAGCGGA 2005
Db 1742 TGATATGGGTGATCGCGATGACAAGCTCGATTCGCAAGCAATGGGGCGTGAGCGGA 1801
QY 2006 AGTGATCTCTCGAATCTGGCAGACTCGGATAGATAAAGAAATTTGGTAAGCTTC 2065
Db 1802 AGTGATCTCTCGAATCTGGCAGACTCGGATAGATAAAGAAATTTGGTAAGCTTC 1861
QY 2066 TGAAGATGGCAAGATGAACATTAATTTCCGATTAAGCGCTACATCTCCAAATACACTAT 2125
Db 1862 TGAAGATGGCGAGATGAACGCAACTTCCGTATCAACGCTACATCTCCAAATACACTAT 1921
QY 2126 CAACCCCGCTTTGACCCACCGCGTGAGCGATATATCGGCTCTGTGTGAAGAGGCAAGAT 2185
Db 1922 TAACCCCGCTCTAACCCATCGGGTAAAGCGAGTATATCGGCTCTGTGTGAAGAGGCAAAAT 1981
QY 2186 CGCGACTTGGTGTGGTAAATCTCTGCTTTTGGCGTAAACCCCAAAATCGTGATCAA 2245
Db 1982 CGCTGATTTGGTGTGGTAAATCTCTGCTTTTGGCGTAAACCCCAAAATCGTGATCAA 2041
QY 2246 AGCGGTATGCTGCTCTCTGAAATGGGCGATTTCTAAACGCTGTGTGCCACTCCGCCA 2305
Db 2042 AGCGGTATGCTGCTCTCTGAAATGGGCGACTCCAACGCTCGTGCTTACACCTCA 2101
QY 2306 ACCGGTTTATTACCGGAAATGTTTGGGCAATCACGCAAGCGCAAAATTTGACACAGCAT 2365
Db 2102 GCCGGTTTATTACCGGAAATGTTTGGGCAATCACGCAAGCGCAAAATTTGACACAGCAT 2161
QY 2366 CACTTTTGTTCMAAGTCGCTATGAATAATGGCGTGAAGAAAGCTGGGCTTAGAGCG 2425
Db 2162 CACTTTTGTTCMAAGTCGCTATGAATAATGGCGTGAAGAAAGCTTAGAGCG 2221
QY 2426 CCAAGTTCTACCGGTCAAAACTGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGA 2485
Db 2222 CAAGGTGCTACCGGTCAAAACTGCGGTAAACATCACCAAGAAAGACTTCAAAATTCACAA 2281
QY 2486 CAAAAAGGCAAAATCAACCGTCTGATCGGAAACCTTCGAGTCTTTGTAGATGGCAAACT 2545
Db 2282 CAAGACGCGCATATCACTGTGATCTCTAAACCTTCGAGTCTTTGTAGATGGCAAACT 2341
QY 2546 CTGACACCTTAACCCACCTCGGAAGTGCCTCTAGCCGAGCGCTACACTTTCTCTAGGC 2605
Db 2342 CTGACACCTTAACCCGCTCTGAAGTGCCTCTAGCCGAGCGCTACACTTTCTCTAGGC 2401

QY 2606 ACAATG 2611
Db 2402 NCAATG 2407
RESULT 8
AX356692
LOCUS
DEFINITION Sequence 10 from Patent EP1176192.
ACCESSION AX356692
VERSION AX356692.1 GI:18674029
KEYWORDS
SOURCE
ORGANISM
Helicobacter felis
Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Helicobacteraceae; Helicobacter.
REFERENCE
1 Kusters,J.G. and Cattoli,G.
Helicobacter felis vaccine
Patent: Ep 1176192-A 10 30-JAN-2002;
Akzo Nobel N.V. (NL)
FEATURES
Location/Qualifiers
1..2407
/organism="Helicobacter felis"
/mol_type="unassigned DNA"
/db_xref="taxon:214"
2..682
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23261.1"
/db_xref="GI:18674030"
/translation="MKLTPEQKFLIYYAGEVARKKRAEGLKLNQPEAIAYISAHIM
DEARRKKTVAELMEECMHFLKDEVMPGVGNMVPDLGVEATFPDGTKLVTWNPWIEP
DEHFKAQVPGCKDIELNAGVETLEVTNEQPKSLIHGSHHFFPANKLKFDE
KAYGRUDIPSGNTRLRIGAGQTRKVKQLIPLOGSKKVIQGNLGVNVIADIRHKHKALEK
AKSHGFIK"
693..2399
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23262.1"
/db_xref="GI:18674031"
/translation="MKMKQEVYNTVYGTGDKVRLGDTDLWAEVEHDDYTTVGEELKF
GAGKTIREGMQSNSPDENTLDLVITNAMIIDYTGIVKADIGIKNGKIHGIGKAGNKD
MODGVSPIVVGTEALAGEMLITAGGIDSHFLSPQOPPTALANGVTTMPGGGT
GPDVGNATITPGKNLHRLRAAEYSNNVFLGKGNSSKKQLVQIEAAGLGPK
LHBDWGTTPSAIDHCLSVADVDVQVCIHDTVNEAGYVDDTLNMGRAIHAYHIEG
AGGSHPDVITMAGELNLPSTPTTPIYTINTVAEHLMLTCHLDRIRREDLOFS
QSRIRPGSIAAEDVLHDIGVIAMTSSDQAMGRAGEVIRPTQWADKNKGFGLPED
SANDNFRIKRYSKYTINPALTHGVSEYIGSVVEGKIADLVWNPAPFGVKPKIVIK
GMVVPFEMSNASVTPQPVYVREMGHHGKAFDTSITFVSKVAVENGVKELGL
ERKPLPVKNCRNITKDFKFNKTAHITVDPKTFEVVDGKLCSTSKPASEVPLAQRIT
FF"
ORIGIN
Query Match 74.9%; Score 2158.6; DB 6; Length 2407;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 155; Indels 0; Gaps 0;
QY 206 GTGAAACTCACACCAAGAGCAAGAAAGTTCTTGTATATATATCGGGCGAAGTGGCT 265
Db 2 GTGAAACTCACACCAAGAGCAAGAAAGTTCTTGTATATATATCGGGCGAAGTGGCT 61
QY 266 AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGGAAGCCATTGCTTACATTAGT 325
Db 62 AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGGAAGCCATTGCTTACATTAGT 121
QY 326 GCCCATATTAGGACGAGCGCGCTGGAAAAAACCCTTGCCAGCTTATGAAGAG 385
Db 122 GCCCATATTAGGACGAGCGCGCTGGAAAAAAGACCGTTGCGGAACCTTATGAAGAG 181

QY	386	TGATGCACTTTTGGAAAAAGATGAAGTAATGCCCCGGGTGGTAAATATATGTTTCCCGAT	445
DB	182	TGTATGCCACTTTTGGAAAAAGACGAGGTGATGCCCGTGTGGGAAATATGTTCCCTGAT	241
QY	446	CTAGGTGTAGAGCCACCTTCTCTGATGTACGAACTTGTAACTGTGAATTTGGCCCATC	505
DB	242	TTAGGCGTGAAGCTACTTTTCCGATGGCAACAACTCGTACCCTGAAATTTGGCCCATC	301
QY	506	GAACCAAGTGAAGCACTTCAAGCGGCGGAAGTGAATTTTGGTTGCGATAAAGACATCGAG	565
DB	302	GAACCCGATGAACACTTCAAGCGGCGGAAGTCAAAATTTGGCTGTGATTAAGACATTTGAA	361
QY	566	CTCAATGCAAGGCAAGAAGTAACCGAACTTTGAGGTTTCTAAATGAAGGGCTTAAATCCTTG	625
DB	362	CTCAACGCAAGTAAAGNAAGTTACCGAACTTAGAAGTTTACCACGAAGCACTTAAATCCTTG	421
QY	626	CATGTGGGTAGCAATTTCCACTTCTTTTGAAGCTTACAGGCCACTAAATTCGATCGTGA	685
DB	422	CATGTGGGTAGCAATTTCCACTTCTTTTGAAGCCAAACAGGCCATTTGAAATTCGATCGGAA	481
QY	686	AAAGCCTATGGCAACGCCCTAGATATTCCTCTGGCAACACGCTACGCAATTTGGGGCAGGA	745
DB	482	AAAGCCTATGGCAACGCCCTAGATATTCCTCTGGCAACACGCTACGCAATTTGGGGCAGGA	541
QY	746	CAAAACCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAAAAGTGAATGGCATGAAC	805
DB	542	CAAAACCGTAAAGTGCAGTTAATCCCTCTTGGCGGCACTTAAAAAAGTGAATGGCATGAAC	601
QY	806	GGGCTTGTGAATAACATCCGGATGAACGCCATTAACATATAAGCGCTTGACAAAGGGCAAA	865
DB	602	GGGCTTGTGAATAATTTGCAGATGAACGCCATTAACACAAAAGCGTTAGAAAAGCAAAA	661
QY	866	TCTCAGCGATTTATCAAGTAAGGAGACTCCCATGAAAAATGAAAAACAAAGATATGTAAA	925
DB	662	TCTCAGCGATTTATCAAAATGAAGAGACTCCCATGAAAAATGAAAAACAAAGATATGTAAA	721
QY	926	TACCTACGGACCCACCAAGGGCGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCAGA	985
DB	722	TACCTACGGACCTACCAAGGCGACAAAGTGGCTTTAGGAGATACCGATCTTTGGGCAGA	781
QY	986	AGTAGAACATGCTATACCACTATGGCGAAGAACTTAAATTTGGCGCGGTAAAACTAT	1045
DB	782	AGTAGAACATGCTATACCACTATGGCGAAGAGCTCAAATTTGGCGCGGTAAAACTAT	841
QY	1046	CGGTGAGGTATGGGTACAGCAATAGCCCTCATGAAACACCCTAGATTTAGTCATCAC	1105
DB	842	CGGTGAGGCGATGGGTACAGCAATAGTCCAGATGAAACACCCTAGATTTAGTCATCAC	901
QY	1106	TAAACGCGATGATTTATCGACTACACCGGGATTTTACAAAGCCGACATTTGGGATTTAAAAACGG	1165
DB	902	CAACGCGATGATTTATGACTACACCGGGATTTTACAAAGCCGACATTTGGGATTTAAAAATGG	961
QY	1166	CAAAATCCATGGCATTTGGCAAGGCAAGAAACAAAGCAATGCAAGATGGCGTAAGCCCTCA	1225
DB	962	CAAAATCCATGGCATTTGGCAAGGCAAGAAACAAAGCAATGCAAGATGGCGTAAGCCCTCA	1021
QY	1226	TATGTGCTGGGTGGGGACAGAGACACTACAGGGAGGTATGATTTATACCGCTGG	1285
DB	1022	TATGTGCTGGGTGGGGACAGAGACATTTAGACGGGGAAGGTATGATTTATACCGCTGG	1081
QY	1286	GGGAATCGAATTCACACCCCACTTCTTTCTCCACAACTTCCCTACCGCTCTAGCCAA	1345
DB	1082	GGGGATCGAATTCACACCCCACTTCTTCTCCACAACTTCCCTACCGCTCTAGCCAA	1141
QY	1346	TGGCGTTACAACCATGTTTGGAGGGCGGACAGGTCTCTGTAGATGGCACGAATTCGACTAC	1405
DB	1142	TGGCGTTACAACCATGTTTGGCGGTGGCACAGGTCCGGTATAGTGGCACGAATTCGACTAC	1201
QY	1406	TATCACTCCGGCAATGGAATTCGACCGGATGTTGGCGGACGACGAAGATTTCTAT	1465
DB	1202	CATCACTCCGGCAATGGAATTCGACCGGATGTTGGCGGACGAGTGAAGATTTCTAT	1261
QY	1466	GAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACTTTGTAGAACAACT	1525

DB	1262	GAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCCAGTTAAAAAACAATTTGTAGAACAAAT	1321
QY	1526	AGAAGCGGGCGGATTTGTTTAAATTTGCAATGAAGACTCTGGGGCACAAACCAAGTGGCAT	1585
DB	1322	AGAAGCGGGCGGATTTGTTTAAATTTGCAATGAAGACTCTGGGGCACAACTCCNAGTGAAT	1381
QY	1586	CGATCACTGCTTTGAGCGTGGCAGATGAATAAGATGTGCAAGTTTGTATTCACACCGGATAC	1645
DB	1382	CGATCACTGCTTTGAGCGTAGCAGATGAATACGATGTGCAAGTTTGTATTCACACCGGATAC	1441
QY	1646	AGTCAATGAGGCGAGGTTATGTAGATGACACCCCTTAAATGCAATGAAAGGGCGGCGCATCCA	1705
DB	1442	GGTCAATGAGGCGAGGTTATGTAGATGACACCCCTTAAATGCAATGAAAGGGCGGCGCATCCA	1501
QY	1706	TGCTTACCACATTTGAGGCGGCTGGAGGACACTCACTCATGTGTTTATCACCATGGCAGG	1765
DB	1502	TGCTTACCACATTTGAGGAGCGGGCGGAGGACACTCACTCATGTGTTTATCACCATGGCAGG	1561
QY	1766	CGAGCTCAATATTTCTTACCCCTCTCTCCACCCACCCCTATTTCCCTATACCAATTAATACGGT	1825
DB	1562	CGAGCTCAATATTTCTTACCCCTCTCTCCACCCACCCCTATTTCCCTATTAATACGGT	1621
QY	1826	TGCAGAACACTTTAGACATGCTCATGACATGCCACCACTTAGACAAACGCAATCCGCGAGGA	1885
DB	1622	TGCAGAACACTTTAGACATGCTCATGACCTGCCACCACTTAGATAAAACGCAATCCGCGAGGA	1681
QY	1886	TTTACAAATTTTCTCAAAGCGGTATCCGCCCGGCTCTATCGCGCTGGAAGATGTGCTCCA	1945
DB	1682	TTTACAAATTTTCCAAAGCGGTATCCGCCCGGCTCTATCGCGCTGGAAGATGTGCTCCA	1741
QY	1946	TGATATGGGTGTGATCGCGATGACAAAGCTCGAATTCGCAAGCAATGGGGCGTGCAGCGGA	2005
DB	1742	TGATATTTGGGTGTGATCGCGATGACAAAGCTCGAATTCGCAAGCAATGGGGCGTGCAGCGGA	1801
QY	2006	AGTGAATTCCTCGAACTTTGGCAGACTCGGATTAAGAAATAAAAAAGAAATTTGGTAAGCTTCC	2065
DB	1802	AGTGAATTCCTCGAACTTTGGCAGACTCGGATTAAGAAATAAAAAAGAAATTTGGTAAGCTTCC	1861
QY	2066	TGAAGATGCAAGATGAAGATTAATTTCCGATTAAGCGCTACATCTCCAATACACTAT	2125
DB	1862	TGAAGATGATGAGATGAAGCAACTTTCCGTATCAAAACGCTACATCTCCAATACACTAT	1921
QY	2126	CAACCCGCTTTGACCCACCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGAT	2185
DB	1922	TAACCCGCTCTTAAACCCATGGGTGAAGCGATATATCGGCTCTGTGGAAGAGGCGCAAAAT	1981
QY	2186	CGCCGACTTTGGTGGTGAATCTTCCCTTTTGGCGTAAACCCCAAAATTCGTGATCAA	2245
DB	1982	CGCTGATTTGGTGGTGGAAATCTTCCCTTTTGGTGTGAACCTTAAGATTTGTGATCAA	2041
QY	2246	AGGCGGTATGGTGGTCTTCTCTGAAATGGGCGATTTCTAAGCGCTCTGTGCGCACTCCCCA	2305
DB	2042	AGGCGGTATGGTGGTCTTCTCTGAAATGGGCGCACTCCAAACGCGTCCGTACACCTCA	2101
QY	2306	ACCGGTTTATACCGGAAATGTTTGGGCACTCACGGCAAGGCGAAATTTGCACACCGCAT	2365
DB	2102	GCGGTTTATTAACCGGAAATGTTTGGGCACTCACGGCAAGGCGAAATTTGCACACCGCAT	2161
QY	2366	CACCTTTGTTTCCAAAGTCGCCCTATGAAATGGCGTGAAGAAAGCTTGGGCTTTAGAGCG	2425
DB	2162	CACCTTTGTTTCCAAAGTCGCCCTATGAAATGGCGTGAAGAAAGCTTAGGCTTAGAGCG	2221
QY	2426	CCAAGTTCTACCGGTCAAAAACTGCCGTAAACATCACCAAGAAAGACTTCAAGTTCAACGA	2485
DB	2222	CAAAGTGTCTACCGGTGAAAAAATTCGCCGCAACATCACTAAGAAAGACTTCAAAATTCACAA	2281
QY	2486	CAAAACGGCAAAATCACCGTCGATCCGAAACCTTTCGAGGTCTTTGTAGATGCACAACT	2545
DB	2282	CAAGACGGCGCATATCACTGTGATCTTAAACCTTTCGAGGTCTTTGTAGATGCACAACT	2341
QY	2546	CTGCACTCTAAACCCACCTCGCAAGTGCCTCTAGCCCGAGCTACACTTTTCTTAGGC	2605

Db	2342	CTGCACCTTAACCCGCTCTGAAGTGGCTTAGCCAGCGCTACACTTTCTTAGGC	2401
Qy	2606	ACAATG 2611	
Db	2402	NCAATG 2407	
RESULT 9			
LOCUS	BD185304	2183 bp	DNA linear PAT 17-JUN-2003
DEFINITION	Helicobacter felis vaccine.		
ACCESSION	BD185304		
VERSION	BD185304.1	GI:31877504	
KEYWORDS	JP 2002355054-A/3.		
SOURCE	Helicobacter felis		
ORGANISM	Helicobacter felis		
REFERENCE	Bacteria: Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.		
AUTHORS	1 (bases 1 to 2183)		
TITLE	Kusters,J.G. and Cattoli,G.		
JOURNAL	Helicobacter felis vaccine		
COMMENT	Patent: JP 2002355054-A 3 10-DEC-2002; AKZO NOBEL NV		
OS	Helicobacter felis		
EN	JP 2002355054-A/3		
PD	10-DEC-2002		
PF	16-JUL-2001 JP 2001214711		
PR	17-JUL-2000 EP 00202565.8		
PI	JOHANNES GERARDUS KUSTERS,GIOVANNI CATTOLI		
PC			
C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC			
A61K39/23,			
PC A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15,			
PC C12N1/19,			
PC			
C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/ PC			
53,			
PC			
G01N33/53,G01N33/566,G01N33/569// (C12N9/80,C12R1:01), (C12Q1/68, PC			
C12R1:01),			
PC C12N15/00,C12N5/00,A61K37/02			
CC Helicobacter felis vaccine			
FH Key	Location/Qualifiers		
FT CDS	(3) . . (683)		
FT CDS	(694) . . (2181).		
FEATURES	Location/Qualifiers		
source	1. .2183		
ORIGIN			
Query Match	67.6%;	Score 1948;	DB 6; Length 2183;
Best Local Similarity	93.3%;	Pred. No. 0;	
Matches 2035; Conservative	0;	Mismatches 145; Indels	0; Gaps 0;
Qy	206	GTGAACCTCACCCCAAGAGCAAGAAAGTCTTGTTATATATGCGGCGAAGTGGCT	265
Db	3	GTGAACCTCACCCCAAGAGCAAGAAAGTCTTGTTATATATGCGGCGAAGTGGCT	62
Qy	266	AGAAAGCGAAAGCAGAGGCGCTTAAAGCTCAACCAACCCGAAGCCATTGCTTACATTAGT	325
Db	63	AGAAAGCGAAAGCAGAGGCTTAAAGCTCAATCAACCCGAAGCCATTGCTTACATTAGT	122
Qy	326	GCCCATATTATGACGAGCGCGCTGGAAAAAACCCTTGCCAGCTTATGGAAGAG	385
Db	123	GCCCATATTATGACGAGCGCGCTGGAAAAAACCCTTGCTGAACCTTATGGAAGAA	182
Qy	386	TGCATGCACCTTTTGAAGAAAGATGAATATGCCGGGTGGGTATATGTTCCCGAT	445
Db	183	TGATGCACCTTTTGAAGAAAGATGAATATGCCGGGTGGGTATATGTTCCCGAT	242
Qy	446	CTAGGTGTAGAAGCCACTTTCCTGATGGTACGAAACTTGTAACTGTGAATTGGCCCATC	505
Db	243	TTGGGCGTGAAGCCCACTTTCCCGATGGCACCAAACTCGTAACCGTGAATTGGCCCAAT	302
Qy	506	GAACCAAGATGAGCACTTCAAGCGGCGAAGTGAATTTGGTTGCGATAAAGACATCGAG	565
Db	303	GAACCTGATGACACTTTAAAGCGGCGTGAATGAATTTGGCTGTGATAAAGACATCGAG	362
Qy	566	CTCAATGCAAGCAAGGTAACCGAATTGAGGTTACTAATGAAGGCGCTTAAATCTCTTG	625
Db	363	CTCAAGCTGGGTAAAGAAAGTTACCGAGCTTGAAGTTACCAACGAAGGACCTTAAATCTCTTG	422
Qy	626	CATGTGGGTAGCCATTTCCACTTTTGAAGCTAAACAAGCACTTAAATTCGATCGTGAA	685
Db	423	CATGTGGGTAGCCATTTCCACTTTTGAAGCAACCAAGGCAATTTGAAATTCGATCGTGAA	482
Qy	686	AAAGCCCTATGGCAAAAGCGCTTAGATATTCCTCTGGCAACACGCTACGCAATTTGGGCGAGA	745
Db	483	AAAGCCCTATGGCAAAAGCGCTTAGATATTCCTCTGGCAACACGCTACGCAATTTGGGCGAGA	542
Qy	746	CAAAACCGCAAGTGCAGTTGATTCCTCTTTGGTGGCAGTAAAAAGTGAATTTGGCATGAAC	805
Db	543	CAAAACCGTAAAGTGCAGTTAAATCCCTCTTTGGCGGTAGTAAAAAGTGAATTTGGCATGAAC	602
Qy	806	GGGCTTGTGAATAACATCGCGGATGAACGCGATAAACATAAAGCGCTTGCAAGCGGCAAA	865
Db	603	GGGCTTGTGAATAATTTGCGGACGAACGCGCATAAACCAAGGCACTAGACAAGGCAAAA	662
Qy	866	TCTCACGGATTTATCAAGTAAAGAGACTCCCATGAAATGAAAAAACAAGAAATATGTAAA	925
Db	663	TCTCACGGATTTATCAAGTAAAGAGACTCCCATGAAATGAAAAAACAAGAGTATGTAAA	722
Qy	926	TACCTACGGACCCCAAGCGGATAAAGTGCCTTAGGAGATACCGATTTTGGGCGAGA	985
Db	723	CACCTACGGACCCCAAGCGGATAAAGTGCCTTAGGAGATACCGATTTTGGGCGAGA	782
Qy	986	AGTAGAACATGACTATACCACTTATGGCGAAGAACTTTAAATTTGGCGGGGTAAAACTAT	1045
Db	783	AGTAGAACATGACTATACCACTTATGGCGAAGAGCTCAAAATTTGGCGGGGTAAAACTAT	842
Qy	1046	CCGTAGGGTATGGTCAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTAGTATCATCAC	1105
Db	843	CCGTAGGGTATGGTCAGAGCAATAGCCAGATGAAAAACACCTTAGATTAGTATCATCAC	902
Qy	1106	TAAACCGCATGATTATGACTACACCGGATTTTCAAAAGCCGACATTTGGGATTAATAACGG	1165
Db	903	CAACCGCATGATTATCGACTACACCGGATTTTAAAGCCGACATTTGTTATTAATAATGG	962
Qy	1166	CAAAATCCATGGCATTTGGCAAGCGAGGAAACAAGGACATGCAAGATGGCGTAAGCCCTCA	1225
Db	963	CAAAATCCATGGCATTTGGCAAGCGGGAACAAAGACATGCAAGATGGCGTAAGCCCTCA	1022
Qy	1226	TATGGTGGTGGTGGGACAGAGCACTAGCAGGGGAGGTATGATTATACCGCTGG	1285
Db	1023	TATGGTGGTGGTGGGACAGAGCACTAGCAGGGGAGGTATGATTATACCGCTGG	1082
Qy	1286	GGGAATCGATTACACACCCCACTTCCCTTCTCCACAACTTCCCTACCGCTCTAGCCAA	1345
Db	1083	GGGAATCGATTGACACACCCCACTTCCCTCTCCCACTTCCCTACCGCTCTAGCCAA	1142
Qy	1346	TGGCGTTAACCAATGTTTGGAGCGGCA CAGGTCCTGTAGATGGCAAGATTCGACTAC	1405
Db	1143	TGGTGTTAACAACCATTTTGGAGGTGGCACAGGTCGCTAGATGGCAAGATTCGACCAC	1202
Qy	1406	TATCACTCCGGGCAATGGAACTTGACCGCATGTTGGCGGACAGAGATTTCTAT	1465
Db	1203	CATCACTCCGGGCAATGGAACTTGACCGCATGTTGGCGGACAGAGATTTCTAT	1262
Qy	1466	GAATGTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTGTAGAACAACT	1525
Db	1263	GAATGTAGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTGTAGAACAACT	1322
Qy	1526	AGAAGCGGCGCATTTGTTTAAATTCATGAAGACTTGGGCGCAACAACCAAGTGCAT	1585

Db 483 AAAGCCTATGGCAAACGCCCTAGATATTCCTCTGGCAAACGCTACGCATTCGGGGCAGGA 542
Qy 746 CAAACCCGCAAAAGTGCAGTTGATTCCTCTTGGGGCAGTAAAAAAGTGATTCGCATGAAC 805
Db 543 CAAACCCGTAAGTGCAGTTAATCCCTCTTGGCGGTAGTAAAAAAGTGATTCGCATGAAC 602
Qy 806 GGGCTTTGTGAATTAACATCGCGGATGAACGGCATAAACATAAAGCGCTTGAACAAGGGGAAA 865
Db 603 GGGCTTTGTGAATTAATTCGGGACGAAACGCATATAACACAAAGACCTAGACAAGGCAAAA 662
Qy 866 TCTCAGCGATTTATCAAGTAAGGAGACTCCCATGAAATAAGTAAATAAACAAGAAATATGTAAA 925
Db 663 TCTCAGCGATTTATCAAGTAAGGAGACTCCCATGAAATAAGTAAATAAACAAGAGTATGTAAA 722
Qy 926 TACCTACGGACCCACCAAGCGGATAAAGTGCCTTTAGGAGATACCGATCTTTGGGCAGA 985
Db 723 CACCTACGGACCCACCAAGCGGATAAAGTGCCTTTAGGAGATACCGATCTTTGGGCAGA 782
Qy 986 AGTAGAACATGACTATACCACTATTCGGCGAAGAACTTAAATTTGGCGCGGGTAAACTAT 1045
Db 783 AGTAGAACATGACTATACCACTATTCGGCGAAGAGCTCAAAATTTGGCGCGGGTAAACTAT 842
Qy 1046 CCGTGAAGGTATGGGTACAGAGCAATAGCCCTGATGAAAAACACCTAGATTTAGTCAAC 1105
Db 843 CCGTGAAGGTATGGGTACAGAGCAATAGCCAGATGAAAAACACCTTAGATTTAGTCAAC 902
Qy 1106 TAAACGGGATGATTTATCACTACACCGGGATTTAAAGCCGACATTTGGGATTTAAAAACGG 1165
Db 903 CAACGGATGATTTATCACTACACCGGGATTTAAAGCCGACATTTGGTATTTAAAAATGG 962
Qy 1166 CAAATTCATGGCATTTGGCAAGGAGGAAACAAGGACATGCAAGATGGCGGTAAAGCCCTCA 1225
Db 963 CAAATTCATGGCATTTGGCAAGGAGGAAACAAGGACATGCAAGATGGCGGTAAAGCCCTCA 1022
Qy 1226 TATGGTCTGGGTGGTGGCAGACAGACCTAGCAGGGAGGTATGATTTATCCGCTGG 1285
Db 1023 TATGGTCTGGGTGGTGGCAGACAGACCTAGCAGGGAGGTATGATTTATCCGCTGG 1082
Qy 1286 GGGAAATCGATTTACACACCCACTTTCCTTCTCCACAACTTCCCTACCGCTCTAGCCAA 1345
Db 1083 GGGAGTGGATTCGACACCCACTTTCCTTCTCCCACTTCCCTACCGCTCTAGCCAA 1142
Qy 1346 TGGCGTTACACCATGTTTGGAGCGGCAACAGTCTCTGTAGTGGCAGCAATCGCACTAC 1405
Db 1143 TGGTGTGTACACCATGTTTGGAGTGGCAGAGTCCGCTAGTGGCAGCAATCGCACTAC 1202
Qy 1406 TATCACTCCGGGCAATAGGAATTCGACCCGATGTTGGCGGCAGCAGAGATTTCTAT 1465
Db 1203 CATCACTCCGGGCAATAGGAATTCGACCCGATGTTGGCGCAGCTGAAGAGTATTTCTAT 1262
Qy 1466 GAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAACAACTTGTAGAACAAAT 1525
Db 1263 GAATGTAGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAACAACTTGTAGAACAAAT 1322
Qy 1526 AGAAGCGGGCGGATTTGGTTTAAATTCATGAGACTGGGGCAACAACCAAGTGGGAT 1585
Db 1323 AGAAGCGGGCGGATTTGGTTTAAATTCATGAGACTGGGGCAACAACCAAGTGGGAT 1382
Qy 1586 CGATCACTGCTTGAGCGTGGCAGATGAATACGATGTCGAAGTTTGTATCCACCGGATAC 1645
Db 1383 CGATCACTGCTTGAGCGTGGCAGATGAATACGATGTCGAAGTTTGTATCCACCGGATAC 1442
Qy 1646 AGTCAATGAGCGAGTTATGATGATGACCCCTAAATGCAATGAACGGCGCGCATCCA 1705
Db 1443 GGTCAATGAGCGAGTTATGATGATGACCCCTAAATGCAATGAACGGCGCGCATCCA 1502
Qy 1706 TGCCTACCACTTGAAGGAGCGGTGGAGGACACTCACCTGATGTTATCAACATGGCAGG 1765
Db 1503 TGCCTACCACTTGAAGGAGCGGTGGAGGACACTCACCTGATGTTATCAACATGGCAGG 1562
Qy 1766 CGAGCTCAATTTCTACCTCTCCACACCCCACTATTCCTCTATACCAATTAACGTT 1825
Db 1563 CGAGCTCAATTTCTACCTCTCCACACCCCACTATTCCTCTATACCAATTAACGTT 1622

Qy 1826 TGCGAAACACTTAGACATGCTCATGATGCCACCACTAGACAAAACGCAATCCCGCAGGA 1885
Db 1623 TGCGAAACACTTAGACATGCTCATGATGCCACCACTAGATAAGCGCATCCCGCAGGA 1682
Qy 1886 TTTTACAAATTTTCTCAAAGCCGTATCCGCCCGGCTCTATCGCGCTCAAGATGTGCTCCA 1945
Db 1683 TTTTACAAATTTTCTCAAAGCCGTATCCGCCCGGCTCTATTTGGCGCTGAGGATGTGCTCCA 1742
Qy 1946 TGATATGGGTGTGATTCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGCTGAGGCGA 2005
Db 1743 TGATATGGGTGTGATTCGCGATGACTAGCTCCGATTCGCAAGCAATGGGGCGCTGAGGCGA 1802
Qy 2006 AGTGATTTCTCGAACTTTGGCAGACTGGCGGATAAGATAAATAAGAAATTTGGTAAGCTTCC 2065
Db 1803 AGTGATTTCTCGAACTTTGGCAAACTGCAGATAAGATAAATAAGAAATTTGGTAAGCTTCC 1862
Qy 2066 TGAAGATGGCAAAAGATAACGATAATTTCCGCATTAAGCGGTACATCTCCAAATACACTAT 2125
Db 1863 TGAAGATGGTGCAGATAACGACAACTTCGCGATCAAAACGCTATATCTCCAAATACACTAT 1922
Qy 2126 CAAACCCCGCTTTGACCCACGCGGTAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGAT 2185
Db 1923 TAATCCGCTTTGACCCATGGCGTGAAGTATATCGGCTCTGTGGAAGAGGCAAGAT 1982
Qy 2186 CGCGACTTGGTGTGGAATCTGCTTTTGGCGTAAAAACCAAAATCGTGATCAA 2245
Db 1983 CGCGACTTGGTGTGGAATCTGCTTTTGGCGTAAAAACCAAAATCGTGATCAA 2042
Qy 2246 AGCGGTATGGTGTCTCTCTGAAATGGCGATTTAAACGCGTCTGTGCCCATCTCCCA 2305
Db 2043 AGTGCGATGGTGTCTCTCTGAAATGGCGATTTAAACGCGTCTGTGCCCATCTCCCA 2102
Qy 2306 ACCGTTTATTTACCGGAAATTTTGGGCTACCGCAAGCGGAAATTTGACACCAAGCAT 2365
Db 2103 GCGGTTTATTTACCGGAAATTTTGGGCTACCGCAAGCGGAAATTTGACACCAAGCAT 2162
Qy 2366 CACTTTTGTTCCTCAAGTCG 2385
Db 2163 CACTTTTGTCTCAAGCG 2182

RESULT 11

AF330621 8406 bp DNA linear BCT 13-MAY-2003
LOCUS Helicobacter bizzeronii tRNA ribosyltransferase gene, partial
DEFINITION cds; GidB, complete cds; urease gene cluster, complete sequence and
unknown gene.
ACCESSION AF330621
VERSION AF330621.1 GI:27462193
KEYWORDS
SOURCE Helicobacter bizzeronii
ORGANISM Helicobacter bizzeronii
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 8406)
AUTHORS Zhu, J., Teng, C.H., Chang, C.F., Chang, C.D., Simpson, K.W., Wei, C.,
McDonough, P., McDonough, S. and Chang, Y.F.
TITLE Cloning and characterization of a Helicobacter bizzeronii urease
gene cluster
JOURNAL DNA Seq. 13 (6), 321-331 (2002)
PUBMED 12652903
REFERENCE 2 (bases 1 to 8406)
AUTHORS Zhu, J. and Chang, Y.F.
TITLE Direct Submission
JOURNAL Submitted (21-DEC-2000) Population Medicine and Diagnostic Science,
College of Veterinary Medicine, Cornell University, Tower Road,
Ithaca, NY 14853, USA
FEATURES
source
1. .8406
/organism="Helicobacter bizzeronii"
/mol_type="genomic DNA"
/db_xref="taxon:56877"

Qy	493	GAATTGGCCCATCGAAACCAAGATGAGACATCTCAAAGCGGGCGAAAGTGAAATTTGGTTGCGGA	552
Db	2628	CCATACCCCGCTTGAAGATGGTGGGCATAAATTTGGCTTCGGGTGAAGTGATTTTGA AAAA	2687
Qy	553	TAAAGACATCGAGCTCAATCAGCGCAAGAGTAAACCGAACCTTTGAGGTTACTAATGAAGG	612
Db	2688	CGAAGACATCACCTTTGAATGAGCGCAAAACGAGCCACCTTTTAAAGATGTCATAACAAAGG	2747
Qy	613	GCCTAAATCTTGTCATGTGGGTAGCCATTTTCCACTTCTTTTGAAGCTTAAACAGGCACCTAAA	672
Db	2748	CGATCGCCCGTGCAAGTGGGCTCCCACTTCCACTTCTTTGAAGTGAATGAAGCTTTTGA	2807
Qy	673	ATTGATCGTGA AAAAGCCATATGCAAAACGCTAGATATTCCTCTGGCAACACGCTACG	732
Db	2808	ATTTGATCGTGA AAAAGCCATATGCAAAACGCTAGACATTTGCTTCTGGAACCGCTGTGCG	2867
Qy	733	CATTGGGGCAGGACAAACCCGCAAAAGTCAGTTGATTTCCCTCTTTGGTGGCAGTAAAAAGT	792
Db	2868	CTTTTGAACCCGGTGAGAAAAAAAACCGTGGAAATTGATTTCAAATTTGGCGGTAAACCAACGCAT	2927
Qy	793	GATTGGGCATCAACCGGCTGTGTAATAACATCGCGGATGAAACGCGCATAAAAATAAAGCGCT	852
Db	2928	TTACGGCTTTAACTCTCTTTGGATCGCGAAGCGGATCTGATGGCAAAAACTTGTCTCT	2987
Qy	853	TGCAAGGGGAAATCTCACGGATTT-----ATCAAG	883
Db	2988	CAAAACGGCCAAAGAACATGGCTTTGGTGTGTAATTTGGCTTGCATAAAAATAAGG	3047
Qy	884	TAAGGAGCTCCCATGAAA---ATGAAAAACAGATATGTAATCTTACGAGCCAC	940
Db	3048	AAAGGCAATCCGATGAAAAAAATCTCTCGAAAGAAATATGTTTCTATGTATGAGCCAC	3107
Qy	941	CAAAAGCGATAAAGTGCCTTTAGGAGATACCGATCTTTGGGCGAAGTGAACATGACTA	1000
Db	3108	TACGGCGATAAAGTGAGATTTGGCGATACGACCTGATCTTAGAAGTCGAANCATGACTG	3167
Qy	1001	TACACCTATGGCGAAGAACTTAAATTTGGCGGGTAAAACTATCCGTGAGGGTATGGG	1060
Db	3168	CACCACTTATGGCGAAGAAATTAAGTTTGGTGGCGGTAAAAACCATTCGCGATGGGATGGC	3227
Qy	1061	TCAGAGCAATAGCCCTGATGAAACACCCTAGATTAGTTCATCTACTAACGGGATGATTAT	1120
Db	3228	ACAAACCAACAGCCCCAGCAGCCACGAATCGATTTGTGCTCACTAACGCCCTGATCGT	3287
Qy	1121	CGACTACACCGGGATTTACAAAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGCGAT	1180
Db	3288	GGATTACACCGGCATTTATAAGCCGATTTTGGCATTTAAAAATGGCAAAATCCATGGCAT	3347
Qy	1181	TGGCAAGCAGGAAACAGGACATCGAAGATGGCGTAAGCCCTCATATGTCGTGGGTGT	1240
Db	3348	TGGCAAAAGCAGGCAATAAAGACATGCAAGATGGCGTTTGGCAACAATCTTTGCGTGGGCC	3407
Qy	1241	GGGCACAGAGCATTAGCAGGGGAAGGATATGATTTATACCGCTGGGGGAATCGATTCA	1300
Db	3408	TGCTACTGAGGCTTTTGGCGCTGAAGGGCTGATTTGTTTACAGCTGGTGGGATTTGACCCCA	3467
Qy	1301	CACCCACTCTCTTCTCCACAAACAATTCCTPACCGCTTAGCCCAATGCGGTTTACAACCAT	1360
Db	3468	CATCACCTTTATTTCTCCCCAACAAATCCCCACAGCATTTTGGCAGCGGATTCACAACCAT	3527
Qy	1361	GTTTGGAGCGGCACAGTCTCTAGATGGCAGAAATGCGACTACTATCACTCCGGGCAA	1420
Db	3528	GATTTGGTGGGGAAACAGTCCAGCTGATGGGACTTAACCGACTACCATCACTCCGGGGCG	3587
Qy	1421	ATGGAACWTGACCGCATTTTGGCGGACAGAGAGATTTCTATGAATGTGGGCTTTTT	1480
Db	3588	CTGGAACCTTTAAAAACATGCTCCGTCGCTCTGAAAGAAATATGCCATGAATTTGGGCTATTT	3647
Qy	1481	GGGCAAAAGGCAATAGCTCTTAGCAAAAAACAATTTGTAGAACAGATGAAGCGGGCGAT	1540
Db	3648	GGGTAAAGGGGAATGTGTCTTATGAACCCCTCCCTGGTCGATCAACTCTGGAAGCTGAGGCCAT	3707
Qy	1541	TGGTTTTTAAATTTGATGAAGACTTGGGGCAACCAACCAAGTCGGATCGATCACTGCTTGAG	1600

RESULT 12

Db	3708	TGGCTTTAAATCCACGAAGACTGGGGTAGCACACTGCAGCGCATCTACCATTTGCTTGAA	3767
Qy	1601	CGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGGATACAGTCAGTAAGTAGGACGAG	1660
Db	3768	TGTGGCTGACAATAAGATGTGCAAGTGGCTATTACACCGGATACCTTGAATGAAGCGGG	3827
Qy	1661	TTATGTAGATGACACCCCTAAATGCAATGAACGGGGCGGCATCCATGCTACACATTTGA	1720
Db	3828	CTGTGTGGAAGACACTTTTGCAAGCCACTGTCTGGGCGCATATCCACACTTTCCACACTGA	3887
Qy	1721	GGGAGCGGTGGAGGACATCTCACTGATGTTATCACCATGGCAGGCGAGCTCAATATTTCT	1780
Db	3888	AGTGTCTGTGTGGGGGACGCTCCGGATGTCAATTAAGATGTCTGGCGAATTTTAACATCCT	3947
Qy	1781	ACCCTCTCCACACACCCCACACTATTCCTCTATACCATTAATACGGTTGCGAGAACACTTAGA	1840
Db	3948	CCGAGCTTCTACCAA CCCACCACTTCTTTCCAGCTGGAATACAGAAGCCGACACATGGA	4007
Qy	1841	CATGCTCATGACATGCCACCACTAGACAAAACGATCCGCGAGGATTTTCAATTTTTCTCA	1900
Db	4008	CATGTTGATGGTGTGCCACCACCTTTGGATAAAAACATCAAGAAGATGTCCAGTTTGTCTGA	4067
Qy	1901	AAGCGTATCCGGCCCCGGCTCTATCCGGCTGAAGATGTCTCCATGATATGGGTGTGAT	1960
Db	4068	TTCTAGATTTGGCCCCCAACCACTGCGCTGAGGACAAACTCCAAGATATGGGATTTT	4127
Qy	1961	CGCGATACAAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGGCGAAGTAGTTCTCTCGAAC	2020
Db	4128	CTCTATCACAGCTCTGACTCCCAAGGGATGGGCCGTGTAGCGGAGTTCATCACCCGCAC	4187
Qy	2021	TTGGCAGACTGGGGATAAGAAATAAAAAGAAATTTGGTGAAGCTTCCTGAAGATGGCAAGA	2080
Db	4188	TTGCAAAACAGCGGACAAAAACAAAAAGAAATTTGGTCTGGCTTGCTGAGGAAAAAGCGCA	4247
Qy	2081	TACGATAAATTTCCGCAATTAAGCGCTACACTCTCCAAATACACTATCAACCCCGCTTTGAC	2140
Db	4248	TAAATGACAACTTCGCGATCAAGCGCTACATTTCCAATACACCATCAACCCCGCTATTGC	4307
Qy	2141	CCACGGCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGATCCCGACCTTGGTGGT	2200
Db	4308	ACACGGCAATTTCTGAATATGTCGGCTCTGTAGAAGTGGGCAAAATTCGCCGATTTGGTGCT	4367
Qy	2201	GTCGAATCTCGCTTTTTTTCGGGTAAACCCAAAATCGTGATCAAAAGCGGATGTTGGTGT	2260
Db	4368	TTGGAGTCTCTCGCTTTCTTTGGCATTAACCCCAACATGATCATCAAAAGCGGATTCATCGC	4427
Qy	2261	CTTCTCTGAATGGCGATTTCTAACGGCTGTGCGCCACTCCCCCAACCGGTTTATTACCG	2320
Db	4428	ACTTTCTCAATATGGCGGATGCCAATGCTCTATCCCCACTCCCCCAACCGGTATTACCG	4487
Qy	2321	CGAAATGTTTGGGCATCACGCGCAGGCGAAATTTGACACCGACATCACTTTTGTTTCCAA	2380
Db	4488	CGAAATGTTTGGCCACCATGGTAAAGCCAAATTTGACACCAATATCACTTTTGTATCCCA	4547
Qy	2381	AGTCGCTATGAANAATCGGTGAAAGAAAGCTGGGCTTAGAGGCGCAAGTTCATCCGGT	2440
Db	4548	AGTGGCTTATGACAAAGCGCATTAAGAAGAGTTGGGCTTGCAGAAAGATGGTTTGGCCAGT	4607
Qy	2441	CARAAACTGGCGTAACTACCAAGAAAGACTTCAAGTTTCAACGACAAAACGCAAAAAT	2500
Db	4608	TAAAAACTGGCGCAACATCACCAAAAAGACCTTCAAAATTCAGCATGTTACCGCACAT	4667
Qy	2501	CACCGTCGATCCGAAAACTTTTCGAGTCTTTGTAGATGGCAAACTCTGCACTCTAAACC	2560
Db	4668	CGAAGTCAATCCTGAACCTTACAAGTTAAGTGGATGGCAAGAGGTTACTTCCRAAGC	4727
Qy	2561	CACCTCGCAAGTGCTCTAGCCACGCGCTACACTTTCTCTAGG	2604
Db	4728	AGCGGATAAATCAGCCTAGCACAACTCTACAACTTGTCTTAG	4771

HECUREASE
LOCUS HECUREASE 2664 bp DNA linear BCT 03-MAY-1995
ACCESSION U000000 Helicobacter heilmannii urease, complete cds's.
VERSION L25079.1 GI:793908
KEYWORDS urease.
SOURCE Candidatus Helicobacter heilmannii
ORGANISM Candidatus Helicobacter heilmannii
REFERENCE Helicobacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
AUTHORS Helicobacteraceae; Helicobacter.
TITLE 1 (bases 1 to 2664)
SOLNICK,J.V., O'Rourke,J., Lee,A. and Tompkins,L.S.
Molecular analysis of urease genes from a newly identified
JOURNAL uncultured species of Helicobacter
PUBMED Infect. Immun. 62 (5), 1631-1638 (1994)
COMMENT 8168924
On May 3, 1995 this sequence version replaced gi:529422.
Original source text: Helicobacter heilmannii (individual_isolate
2) DNA.

FEATURES
 source Location/Qualifiers
 1..2664
 /organism="Candidatus Helicobacter heilmannii"
 /mol_type="genomic DNA"
 /isolate="2"
 /db_xref="taxon:35817"
 /lab_host="BALB/c"
 200..204
 211..915
 /codon_start=1
 /transl_table=11
 /product="urease"
 /protein_id="AAA65722.1"
 /db_xref="GI:529423"
 /tranlation="MKLTPEKLDKMLHVAGELAKORKAKGIKLVNYTEAVALISAHVM
 EARACKSVADLMQGRILLKADDVMPGVAMHIVHGVEAGFPDGTKLVTHTPYEA
 GSDKLAPGVILNKEDNTINAGHKAHLVKNGKDRPVQVGHSHFVFENKLLDFORE
 KAYKRLDIASGTAVRPFEPGEKTVELIDIGGNKRIYGFNALVDROADHDGKKLALUR
 AKEKHPTINGCGDNKR"
 RBS 919..923
 CDS 930..2636
 /codon_start=1
 /transl_table=11
 /product="urease"
 /protein_id="AAA65723.1"
 /db_xref="GI:529424"
 /tranlation="MKKISRKEVSMYGPTTGDKVRLGDTDLILEVEHDCTTYGEEK
 FGGKTRIDMGQTNTPSSSHELPLVITNALI VDYTGIIYKADI GKNGKIHGIGKAGN
 TLQPDGNCNLTCVPATLEALAEGLIVTAGGIDTHIFIPQIPTAFASGITTMGGG
 DGADGYNATTTPGRWNLKEMLEEYAMNLGYLKGNVSPFEPAIDQLSEAGTGF
 KIHEDGSTPSIAKNHALTKDVQVAJHTDTLNAGCVEDTLEAIAGRTTHTTFEFA
 GAGGHAPDIVKWAGEFNILPASTNPITIPTKNTAEHMDMLWCHLDKNIKEDVEFA
 DSRIRPQTAAEDKLDHMGIFSITSDDSQAMRGVETITRWQTADKNKKEPGRLPEE
 KGDNDNFRIKYSKITYPATHIGISEYSGVEVGKYADLVLSPPAFFGIKPNNLIK
 GGFIALSQGDANASITPPQPVYREMFHHGKAKEDTNIITFVSQVAYENGIKHLEJL
 QRVLVPVKCNRLITKKDLKFENDVTAHIEVNPETYKYKVDGNEVTSHAADKLSLAQLYN
 Lf"

ORIGIN

Query Match 39.5%; Score 1138; DB 1; Length 2664;
Best Local Similarity 66.4%; Pred. No. 6.8e-217;
Matches 1709; Conservative 1; Mismatches 826; Indels 36; Gaps 4;

QY 63 TTGTTAATRCTATTATTTTAAATAACTTACTTATCATATATATAATATATT 122
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 TTTGTAATGAAGTATCAAGAATAAATAAT--ATTATCTTTGGATAACAAATAATTAAC 127
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 123 ACTTATATAAAAAGTTAATAAAAAAGTAAACGAAATTTAGGACTATAATCCCATTCGCCTTTA 182
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 128 AACCATTAGCCAAAAGATACTAAAACTCGCTCGTTGATAGCTTGGCTATTCAATAACA 187
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 183 AAATTAAACACAGGAGTAAATAGGTGAACACTCACACCCAAAGACGAAGAAAGTCTTTGT 242
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 188 CCATTTTATACGAGGAGTTTAGATCAAACTGACACTAAAGAGTTGGTAAGTTCGATGC 247
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1292 CGATTACACACCCACTTCTCTTCTCCACAACTTCCCTACCGCTCTAGCCAAATGGCGT 1351
Db 1328 CGACACCCACATCCACTTCATTCTCTCTCAACAAATCCCTACTGCTTTTGGTAGCGGAAT 1387
QY 1352 TACAACCATGTTTGGAGGGGGACAGCTCCTGTAGATGCGCAGAAATGCGACTACTATCAC 1411
Db 1388 CACCACCATGATCGGGGGGGGACTGGAATCTGCAGATGCGACCAACCGCCACCAATCAC 1447
QY 1412 TCGGGCAAAATGGAATCTGACCGCATGTTGGCGCGAGAGAGATATTTCTATGAATGT 1471
Db 1448 TCCGGCGCGCTGGAACTTTAAAGAAATGCTCCGCGCTTCTGAAGAAATACGCCATGAACCT 1507
QY 1472 GGGCTTTTGGGCAAAAGGCATAGCTCTAGCAAAAACAACTTGTGTAGAACAGTAGAAGC 1531
Db 1508 TGCGTACCTTGGTAAGGGGAATGTCTTTTGAACCTGCTCTCATTTGACCAAGCTAGAAGC 1567
QY 1532 GGGCGCGATTGGTTTAAATTTGATGAAGACTGGGGCACAACACCAAGTGCAGATCGATCA 1591
Db 1568 CGGGCGGATTGGTTTAAATTTCAAGAGACTGGGGCAGCACACCTTCAGCCATCAACCA 1627
QY 1592 CTGCTTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACCGATACAGTCAA 1651
Db 1628 CGCTCTAAACATCGCTGACAAATACGATGTGCAAGTGGCGATCCACACGACACCTTGAA 1687
QY 1652 TGAGGCAAGTTATGTAGATGACACCTTAATGCAATGAACGGCGCGCCATCATGCTTA 1711
Db 1688 TGAAGCGCGTGTGTGGAAGACACCTTAGAAGCCATCGCTGGACGACCACTCCACACCTT 1747
QY 1712 CCATTTGAGGAGCGGTGGAGACACTCACCTGATGTTATCACCATGGCAGCGAGCT 1771
Db 1748 CACACCGAAGTGTGTGGCGGAGACACGCTCTGTACGTGATCAAAATGGCGGCGAAT 1807
QY 1772 CAATATTTACCCCTCTCCACACCCCACTATTTCCCTTATPACCATTAATACGGTTGCGA 1831
Db 1808 TAACATCTTCTGCTTCTTACCAACCCCACTCCCTTTTCCACCAAAACACAGAAGCGA 1867
QY 1832 ACATTTAGACATGCTATGATGACATGCGACACACTAGACAAACGATCGCGAGATTTACA 1891
Db 1868 ACACATGGACATGTTATG---TGCCACCACTTTGGATAAACCAATCAAGAAAGTTGGA 1924
QY 1892 ATTTTCTCAAAGCGTATCGCCCGCTCTATCGGGCTGGAAGATGTGCTCCATGATAT 1951
Db 1925 ATTGTGCGACTACGATATCGCCCGCCCAACCATTTGGCGGGAAGATTAATCTCCACGACAT 1984
QY 1952 GGGTGTGATCGGATGACAAAGCTCGGATTCGCAAGCAATGGGCGGTGCGAGCGCAAGTGAT 2011
Db 1985 GGGGATTTTCTCATCACTAGCTCGACTCTCAAGCCATGGGTGCGTGGCGGAGTGAT 2044
QY 2012 TCCTGAACTTTGGCAGACTCGGATAGAAATAAAGAAATTTGGTAAGCTTCTCTGAAGA 2071
Db 2045 CACCCGCACTTGGCAACAGCGGACAAAAACAAAAAGAAATTTGGCGGCTTGGCTGAAGA 2104
QY 2072 TGCAAGATAACGATTAATTTCCGATTAAGCGCTACATCTCCAAATACATATCAACCC 2131
Db 2105 AAAAGGCGACAAAGCAACTTCGCGATCAAAACGCTATATTTCCAAATACACCATCAACCC 2164
QY 2132 CGCTTTGACCCACCGCGTGAAGGATATATCGGCTCTGTGGAAGAGGCGCAAGATCGCGCA 2191
Db 2165 TGCCATCACACACGGCATTTCTGNATATGTGCGCTCTGTAGAGTGGGTAAATACGCTGA 2224
QY 2192 CTTGGTGGTGTGAATCCTGCTTTTGGCGTAAACACCAAAATCGTGTATCAAGGCGG 2251
Db 2225 CTTGGTGGTGTGGAGCCCTGCTTCTTTGGCAATTAACACCAACATGATCATCAAGGCGG 2284
QY 2252 TATGGTGGTCTTCTGAAATGGCGATTTCTAAGCGGTCTGTGCCCACTCCCAACCGGT 2311
Db 2285 TTTTCATTCGCGCTTTCTCAAAATGGCGATGCGCAACGCTTCTATCCCACTCCCAACCGGT 2344
QY 2312 TTATTTACCGGAATGTTTGGGCATCAGCGCAAGCGGCAAAATTTGACACACGACATCACTTT 2371
Db 2345 GTACTACCGGAATGTTTGGCCACCAAGCGCAAGCGCAAAATTTGACACCAACATCACTTT 2404
QY 2372 TGTTTCCAAAGTCGCTATGAAATATGGCGTGAAGAAAGCTGGGCTTTAGAGCGCCAAAGT 2431

Db 2405 CGGTCTCAAGTGGCTTATGAAAACGGCATTAACACGAGTTGGCTTGCAAAAGTTGT 2464
QY 2432 TCTACCGGTCAAAAACATGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGCAAAAC 2491
Db 2465 GTTGCCTGTGAAAACATGCGCGCAACATCACCAAAAAGACCTTAAGTTCAACGATGTAC 2524
QY 2492 GGCAAAATCACCGTGCATCGGAAAACCTTCGAGGTCTTTGTAGATGCGAACTCTGCAC 2551
Db 2525 CGCACACATCGAAGTCAACCTGAAACCTCAAAAGTGAAGTGGATGCGAACGAGTTAC 2584
QY 2552 CTCTAAACCCACCTCGCAAGTGCCTCTAGCCAGCGCTACACTTTCTTTCTAG 2603
Db 2585 CTCCCATGCGGCTGACAAATTTGAGCCCTAGCACAACTCTATACCTATTCTAG 2636

RESULT 13
AE001446/c 12037 bp DNA linear BCT 20-JAN-1999
DEFINITION Helicobacter pylori, strain J99 section 7 of 132 of the complete genome.
ACCESSION AE001446
VERSION AE001446.1 GI:4154573
KEYWORDS Helicobacter pylori J99
SOURCE Helicobacter pylori J99
ORGANISM Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 12037)
AUTHORS Alm, R.A., Ling, L.S., Moir, D.T., King, B.L., Brown, E.D., Doig, P.C., Smith, D.R., Noonan, B., Guild, B.C., deJonge, B.L., Carmel, G., Tummino, P.J., Caruso, A., Uria-Nickelsen, M., Mills, D.M., Ives, C., Gibson, R., Merberg, D., Mills, S.D., Jiang, Q., Taylor, D.E., Vovis, G.F. and Trust, T.J.
TITLE Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori
JOURNAL Nature 397 (6715), 176-180 (1999)
PUBMED 9923682
REFERENCE 2 (bases 1 to 12037)
AUTHORS King, B.L., Alm, R.A. and Trust, T.J.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-1999) Astra Research Center Boston, 128 Sidney Street, Cambridge, MA 02139, USA
COMMENT Address all correspondence to: hp@arcb.us.astra.com or Richard A. Alm, Astra Research Center Boston, 128 Sidney Street, Cambridge, MA, 02139. Lo-See L. Ling, Donald T. Moir, Douglas R. Smith, Braydon C. Guild, Gilles Carmel, Anthony Caruso, Debra M. Mills, Rene Gibson, and Gerald F. Vovis are with Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA, 02453. Qin Jiang and Diane E. Taylor are with the University of Alberta Department of Medical Microbiology and Immunology, Edmonton, Alberta, Canada, T6G 2H7 and the Canadian Bacterial Diseases Network. All other authors are with Astra Research Center Boston, 128 Sidney Street, Cambridge, MA, 02139. Putative identifications, sequence alignments, and name and sequence search capability are available at ARCB's World Wide Web site. (URL: http://www.astra-boston.com/hpylori).
FEATURES
source Location/Qualifiers
1..12037
/organism="Helicobacter pylori J99"
/mol_type="genomic DNA"
/strain="J99"
/db_xref="taxon:85963"
complement(143..1852)
/gene="ureB"
/note="synonym: jhp0067"
complement(143..1852)
/gene="ureB"
/note="similar to H. pylori 26695 gene HP0072"
/codon_start=1
/transl_table=11
/product="UREASE BETA SUBUNIT"
/protein_id="AAD05651.1"

Query Match	37.8%;	Score 1090.8;	DB 1;	Length 12037;		
Best Local Similarity	65.3%;	Pred. No. 1.4e-207;				
Matches 1681;	Conservative 1;	Mismatches 853;	Indels 39;	Gaps 4;		
Qy	67	TAATRCCTATTAAATTTTAAATAATTAACCTT-ATTATCATATATAATAATATATTA	125	1695	AAGGATGAGCCAAATCCAAACAACCCCTAGCAAAAGAACTGGATTTGATCATCACTAAACG	
Db	2712	TAATACCTAAATTCATTTTAAATAATAATTAAGTAAATAGTAAGCGCTTCTGTTAATCTTAGTAAA	2653	1111	CGATGATTATCGACTACACCGGGATTTTCAAAAGCCGACATTTGGGATTTAAAAACGGCAAAA	
Qy	126	TATATTTAAAAAGTTAATAAAAAAGTAAACGAAATTTAGGACTATATATCCCATTTGCCCTTTAAAAA	185	1635	CTTTAATCGTGGATTACACCGGTATTTATAAGCGGATATTTGGTATTTAAAGACGGCAGAAA	
Db	2652	TCAAAAGATTGCTACAAATTACTCCACCTTGATTGCGTTATGTCCTTCAAGGAAAAAACAC	2593	1171	TCCATGGCATTTGGCAAGCGAGAAACAAGGACATGCAAGATGGCGTGTAAAAACAATCTTA	
Qy	186	TTTAAACAAAGGAGTAAATAGTGAATCTACACCCAAAGAGCAAGAAAGTTCTTGTTAT	245	1575	TCGCTGGCATTTGGTAAAGGCGGTAAACAAAGACATGCAAGATGGCGTGTAAAAACAATCTTA	
Db	2592	TTTTAAGAAATAGGAGAAATGAGATGAATCACTACCCCAAAGAGTTAGATAAAGTTGATGCTCC	2533	1231	TCGTGGGTGGGACACAGACACTAGCAGGGGAAAGGTATGATTATTAACGCTGGGGGAA	
Qy	246	ATTATGGGCGAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCG	305	1515	GGGTGGGTCTCTACTAGAACGCGCTAGCTGGTGAAGCGCTTAATTCGTAACGGCTGGGTGA	
Db	2532	ACTACGCTGGAGAAATTAGCTAGGAAACGCAAAAGAAAGGACATTAAGCTTAACTATGTGG	2473	1291	TCGATTTACACACCACTTCTCTTCTCCACAACAATTCCTACCGCTCTAGCCAAATGGCG	
Qy	306	AAGCCATTGCTTACATTTAGTGCCCATATTTATGACGAAGCGCGCGTGGAAAAAAAACCG	365	1455	TTGACACACATCCACTTCATTTTACCCCAACAATCCCTACAGCTTTTGGCAAGCGGTG	
Db	2472	AAGCGGTAGCTTTGATTAGTGCCCATATTTATGAGAGGAGAGCTGGTAAAAAGACTG	2413	1351	TTTAAACCAATGTTTGGAGCGGCACAGGTCCTGTAGTAGGCAAGATGCGACTACTATCA	
Qy	366	TTGCCAGCTTATGGAAGAGTGCATGCACTTTTGTAAAAAAGATGAAGTAATGCCCGGG	425	1395	TAAACAACCATGATTTGGTGGCGAAACCGGTCTCTGATGGCACTAATTCGACTACTATCA	
Db	2412	CGGCTGAATTTGATGCAAGAGGGCGCACTCTTTTAAAAACCGATGATGTGATGGATGGTG	2353	1411	CTCCGGGCAAAATGGAACCTTTGCAACCGCATTTGCGCGCAGCAGAGAGTATTTCTATGAATG	
Qy	426	TGGGTAATATGGTTCCCGATCTAGGTGTAGAAGCCACTTTCTGATGGTACGAACCTTG	485	1335	CTCCAGGTAGAGAAAATCTTAAAAATGGATGCTCAGAGCGGCTGAAGAATATTTCTATGAAC	
Db	2352	TGGCAAGCATGATCCATGAAATGGGTATTTAAGCGATGTTTCTGATGGGACCAAACTCG	2293	1471	TGGGCTTTTTTGGCAAAAGGCAATAGCTCTAGCAAAAAAACAATTTGTAGAAACAAGTAGAAG	
Qy	486	TAACTGTGAATTTGGCCCATCGAAACCATGAGCACTTCAAAGCGGGCGAAGTGAATTTTG	545	1275	TAGGTTTTCTTAGTAAAGGTAAAGCTTTCTAAACGACGGAGCTTAGCGGATCAAAATGAAG	
Db	2292	TAAACCTGCATACCCCTTTTGGGCTTAATGCTAAATTTGGTTCTGTTGAGTTG---TTC	2236	1531	CGGCGCGCATTTGGTTTTAAATTTGCATGAAGACTTGGGGCAACAACAAGTGCAGTCGATC	
Qy	546	GTTCGATAAAGACATCGAGCTCAATGACGGCAAGAAAGTAAACCGAACTTTGAGTTACTA	605	1215	CGGTCGCAATTTGGCTTTTAAAAATCCACGAAGACTTGGGGCACCACTCTCTTCGCAATCAATC	
Db	2235	TAAAAAATGAAGACATCACTATCAACGAAGCAAAAAAAGCCGTAGCGTGAAGTTAAAA	2176	1591	ACTGCTTGAGCTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACGATACAGTCA	
Qy	606	ATGAAGGCCCTAAATCTTCGATGTGGTAGCCATTTTCCACTTTCTTTGAAGCTTAAACAGG	665	1155	ATGCGTTAGACGTTGCGGCAAAATACGATGTGCAAGTCGCTATTCACACAGACACTTTGA	
Db	2175	ACGTGGGCGACAGACCGGTTCAAATCGGTTTCAACACTTTCCATTTCTTTGAAGTGAATAGAT	2116	1651	ATGAGGACGTTATGTAGATGACACCTTAAATGCAATGAACGGCGCGCCATCCATGCCT	
Qy	666	CACATAAATTCGATCGTGAAAAAGCCCTATGCGCAACCGCTAGATATTCCTCTGGCAACA	725	1095	ATGAAGCCGTTGCTGTTGGAAGACACTATGGCAGCTATTTGCCGAGCAGCTATGACACATT	
Db	2115	GCCTAGACTTTGACAGAGAAAAAACTTTGCGCAACCGCTTAGACATTTGCGAGCGGACAG	2056	1711	ACCACATTGAGGAGCGGGTGGAGGACACTCACCTGATGTTATCACCATCGCAGGCGAGC	
Qy	726	CGCTACGCTATGGGCGAGGACAAACCCGCAAGTGCAGTTGATTCCTTTGGTGGCAGTA	785	1035	TCCACACTGAAGCGCTGGTGGCGGACACGCTCCTGACATTTATAAAGTGGCGGTGAAC	
Db	2055	CGGTAAAGTTTGAAGCTTGGCGGCAAAAAATCCCGTGAAGATTTGATGACATTTGGTGGCAACA	1996	1771	TCAATATTTACCTCTCCACACCCCACTATTCCTATACCATTAATACGTTTCAGTTCGAG	
Qy	786	AAAAAGTGATTGGCATGAAACGGGCTTGTGAATAACATCGCGGATGAACGCCATAACATA	845	975	ACAACATCTCTACCCGCTTCCACTTAACCCCACTATCCCTTTCTGTTGAATACAGAACCG	
Db	1995	GAAGAAATCTTTGGATTTTAAACGATTTGATAGGCAAGCAGACAAACGAAAGCAAAAAA	1936	1831	AACACTTTAGACATGCTCATGACATGCCACCTAGACAAACGATCCCGAGGATTTAC	
Qy	846	AAGCGCTTGACAGGCGAAATCTCACCGATTT-----	877	915	AACATAGGACATGCTCATGTTGTCACACCTTTGATATAAGCATTTAAAGAGATGTCC	
Db	1935	TTGCTTTTACACAGAGCTTAAGAGCGTGGTTTTTCAATGCTGTTAAAGCGATGACAACTATG	1876	1891	AATTTTTCTCAAAGCCGCTATCCGCCCGGCTCTATCCGGCTGAAAGATGTGCTCCATGATA	
Qy	878	----ATCAAGTAAGGAGACTCCCATGAAA---ATGAAAAAACAGAAATATGTAATACCT	930	855	AGTTTCGCTGATTCAGGATCCGCTCAAAACCATTTGCGGCTGAAGACACTTTTGCATGACA	
Db	1875	TAAAAACAAATTAAGGATTAAGAAATGAAGAAATAGCAGAAAAAGATATGTTTCTATGT	1816	1951	TGGGTGTATCGGATGACAAAGCTTCGGATTCGCAAGCAATTTGGGCGTGGCAGGAGTGA	
Qy	931	ACGACCCCAACAAAGGCGATAAAGTGGCTTAGAGATACCGATCTTTTGGGCGAGAGTAG	990	795	TGGGGATTTTCTCAATCACCAGTTCTGACTCTCAAGCGATGGGTGCTGTGGGTGAAGTTA	
Db	1815	ATGGCCCTACACAGGGGATAAAGTGAATTTGGGCGATACAGATTTTGATCGCTGAAGTAG	1756	2011	TTCTCTCGAATTTGGCAGACTCGCGGATAAGAAATAAAAAAGAAATTTTGGTAAAGTCTTCAAG	
Qy	991	AACATGACTATACCACCTATGCGCAAGAACTTAAATTTGGCGCGGTAAACATATCCGTG	1050	735	TCATAGGACTTTGGCAAAACAGCTGACAAAAACAAAAAAGAAATTTTGGCCGCTTGAAGAG	
Db	1755	AACATGACTACACCATTTATGGCAAGAGCTTTAAATTTGGTGGCGGTAAACCCCTAAGAG	1696	2071	ATGGCAAGATAACGATAAATTTCCGCAATTAAGCGCTACATCTCCAAAATACATATCAACC	
Qy	1051	AGGTAATGGGTGACAGCAATAGCCCTGATGAAAAACACCCCTAGATTTTATGCTCATCACTAACG	1110	675	AAAAAGCGATACGACAACTTCAGGATCAACCGTACTTTGCTCTAAATACACCAATTAACC	
Db				2131	CCGCTTTGACCCACGGCGTGAGCGAGTATATTCGGCTCTGTGGAAAGAGGGGCAAGATTCGGCG	
Qy				2190		

FEATURES		Location/Qualifiers
source	1. .10903	/organism="Helicobacter pylori 26695"
		/mol_type="genomic DNA"
		/strain="26695"
		/db_xref="taxon:85962"
gene		/complement(127. .1836)
CDS		/gene="HP0072"
		/complement(127. .1836)
		/note="similar to GB:M60398 GB:X17079 SP:PI4917 PID:149011
		GB:AE000511 percent identity: 100.00; identified by
		sequence similarity; putative"
		/codon_start=1
		/transl_table=11
		/product="urease beta subunit (urea amidohydrolase)
		(ureB)"
		/protein_id="AAD07143.1"
		/db_xref="GI:2313153"
		/translation="MKKISRKEYSMYGGTTGDKVRLGDTLIAEVEHDYTYIGELK
		FGGCTREGMSQSNPKSEKELDLIITNALIVDYTGIVKADIGIKDKIAGIKGKGN
		DMQCVKNNLSVGPALVAGEGLIVTAGDIGLTHIFISPOQIPTAFASGVTTMIGGG
		TGPADGTWATITIPGRNLKMWLRAREEYSNNLGPLAKGNASNDASLADQIEBAGTGF
		KIHEDWGTPDIIKVAAGENILPASTNPTIPTFTVNTAEHMDMLVCHHLDKSIKEDVQF
		GAGGHAAPDIIKVAAGENILPASTNPTIPTFTVNTAEHMDMLVCHHLDKSIKEDVQF
		ADSRIRPTIAAEDLTGDMGIFSISSDSQAMRGVGEVITRTWTADKNKBFKRLKI
		EKGDNDFRIKRYLSKYITINPAIHGISEYVGSVEGVKVDLVLSPAFGVKPNMII
		KGGFTALSQMGDANASIPPTOPVYVREMFHHGKAKYDANITFVSQAAVDKGIKBEGL
		LERQVLPVKNCRNITKDMQFNDTTAHIEVNPTYHVPVDGKEVTSKPANKVSLAQLF
		SIF"
gene		/complement(1840. .2556)
CDS		/gene="HP0073"
		/complement(1840. .2556)
		/gene="HP0073"
		/note="similar to SP:PI4916 PID:149010 PID:43634
		GB:AE000511 PID:1567116 percent identity: 100.00;
		identified by sequence similarity; putative"
		/codon_start=1
		/transl_table=11
		/product="urease alpha subunit (ureA) (urea
		amidohydrolase)"
		/protein_id="AAD07144.1"
		/db_xref="GI:2313154"
		/translation="MKLTPEKLDLMLHYAGELAKKRKKGKIKLVYEAVALISAHIM
		BEARGKKTAAELMQEGRITLLKPDVMDGVASMIHEVGTEAMFPDGTKLVTVHTPTIEA
		NGKLVGPELFLKNEDITINEGKVASVKVKNVGRDPRVQIGSHFHFPEVNRCLDPDREK
		TKGKLDIASGTAVRPFEPCEBKSEVELIDIGNNRIFGPNALVDROADNESKIALHRA
gene		/complement(2809. .2881)
tRNA		/gene="CRNA-Val-2"
		/complement(2809. .2881)
		/gene="CRNA-Val-2"
		/product="crRNA-Val"
gene		/note="codon recognized: GUC"
		/complement(2902. .3375)
CDS		/gene="HP0074"
		/complement(2902. .3375)
		/gene="HP0074"
		/note="similar to GB:M60398 SP:P25178 PID:149009
		GB:AE000511 PID:2313155 percent identity: 100.00;
		identified by sequence similarity; putative"
		/codon_start=1
		/transl_table=11
		/product="lipoprotein signal peptidase (lspA)"
		/protein_id="AAD07145.1"
		/db_xref="GI:2313155"
		/translation="MLKTTKKSLLVFMGVFFLIFGVDOAKYVAILLEGFRYESLMIDIV
		LVFNKGVAFLSLSGEGLKYLOILLGLFLMRQRELKFNKHAIEFGMVFCAVSN
		VLDKPVHGGVVDVYVYHYGFDFAIFNFADVMDVGVGLLLKQFFPFKQKNKKA"
gene		/complement(3369. .4706)
CDS		/gene="HP0075"
		/complement(3369. .4706)
		/gene="HP0075"

615	CACGGATCGCTCATGGGATTACGACGATATGTAGGTCTCTGTAGAAGTGGCGCAAAAGTGGCTG	556
2191	ACTTGGTGGTGGGAATCCTCGCCTTTTTTGGCGTAAAAACCCAAAATCGTGATCAAAAGCGG	2250
555	ACTTGGTATTGTGGAGTCCAGCATCTTTTGGCGTGAACCCCAACATGATCATCAAGGCG	496
2251	GTATGGTGGTCTTCTGTAATGGGCGATTCTTAACGGGTCTGTGCCACTCCCCAAACCGG	2310
495	GATTCTATTGGTTAAAGCCAAATGGGCGATCGCAACGCTTCTATCCCTACCCCAACCGG	436
2311	TTTATTACCGCGAAATGTTTGGGCATCAGCGAAGCGGAAATTTTGACACAGATCACATTT	2370
435	TTTATTACAGAAATGTTTCGCTCACATGTAAGCTTAATACGATGAACATCACATTT	376
2371	TTGTTTCCAAAGTCGCTATGAAATGGGCTGAAAGAAAGCTGGGCTTAGAGCGCCAAG	2430
375	TTGTGCTCAGCGGCTTATGACAAAGCATTAAGAAGATTAGGGCTTGAAGACAAG	316
2431	TTCTACCGGTCAAAACTCGCGTAACATCACCAAGAAAGACTTCAAGTTCAACGACAAA	2490
315	TGTTGCGGTTAAAAATTGCAGAAACATCACTAAAAAAGACATGCAATTCATGACACTA	256
2491	CGGCAAAATCACCGTCGATCCGAAACCTTCGAGGCTCTTGTAGATGCAAACTCTGCA	2550
255	CCGCTCACATTTGAAGTCAATCTGAAACCTTACCATGTTCGTGGATGCGAAAGAAGTCA	196
2551	CCCTAAACCCACTCGCAAGTGCCCTCTACGCCAGCGCTACACTTTCTTCTAGG	2604
195	CTTCTAAACAGTAATAAGTGAGCTTGCTCAACTCTTTAGCATTTTCTAGG	142

RESULT 14		
AE000529/c		
LOCUS	10903 bp	DNA linear BCT 06-APR-1999
DEFINITION	Helicobacter pylori 26695 section 7 of 134 of the complete genome.	
ACCESSION	AE000529	AE000511
VERSION	AE000529.1	GI:2313152
KEYWORDS		
SOURCE	Helicobacter pylori 26695	
ORGANISM	Helicobacter pylori 26695	
REFERENCE	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.	
AUTHORS	1 (bases 1 to 10903)	
	Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G.,	
	Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S.,	
	Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F.,	
	Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G.,	
	Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D.,	
	Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R.,	
	Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and	
	Venter, J.C.	
TITLE	The complete genome sequence of the gastric pathogen Helicobacter	
JOURNAL	Pylori	
PUBLISHED	Nature 388 (6642), 539-547 (1997)	
REFERENCE	9252185	
AUTHORS	2 (bases 1 to 10903)	
	Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G.,	
	Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S.,	
	Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F.,	
	Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G.,	
	Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D.,	
	Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R.,	
	Peterson, J.D., Kelley, J.M., Cotton, M.D., Weidman, J.M., Fujii, C.,	
	Bowman, C., Wathey, L., Wallin, B., Hayes, W.S., Borodovsky, M.,	
	Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C.	
TITLE	Direct Submission	
JOURNAL	Submitted (06-AUG-1997) The Institute for Genomic Research, 9712	
	Medical Center Dr. Rockville, MD 20850, USA	
REFERENCE	3 (bases 1 to 10903)	
AUTHORS	White, O.	
TITLE	Direct Submission	
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712	
	Medical Center Dr. Rockville, MD 20850, USA	

Db 1803 ATGTATGGCCCTACTACAGCGGATNAAGTGAAGATTGGGGGATACAGACTTTGATCGCTGAA 1744
QY 987 GTAGAACTGACTATATCCACTATGCGGAAGAACTTTAAATTTTGGCGGGTAAAACTATC 1046
Db 1743 GTAGAACATGACTACACCATTTATGCGGAAGAGCTTAAATTCGGTGGCGGTAAACCCCTG 1684
QY 1047 CGTGAAGGTATGGGTGAGAGCAATAGCCCTGATGAAACACCCCTAGATTGATGTCATCT 1106
Db 1683 AGAAGAGGATGAGCCAAATCCAACACCCCTAGCAAGAGAAATTGGATCTAATCATCACT 1624
QY 1107 AACGCGATGATTATCGACTACACCGGGATTACAAAGCCGACATTCGGGATTAAGACGCG 1166
Db 1623 AACGCTTTAATCGTGGATTACACCGGTATTTTAAAGCGGATATGGTATTAAGATGCG 1564
QY 1167 AAAATCCATGGGATTCGCAAGGAGGAAACAAGGACATGCAAGATGGCGTAAGCCCTCAT 1226
Db 1563 AAAATCGCTGGCATTTGGTAAAGCGGTAAACAGACATGCAGATGGCGTTAAACAAAT 1504
QY 1227 ATGGTCGTGGGTGTTGGGCAAGAGCACTAGCAAGGGAAAGGTATGATTAATACCGCTGGG 1286
Db 1503 CTTAGCGTAGGTCCTGCTACTGAAGCCTTAGCGGTGAAGGTTTGATCGTAACTGCTGGT 1444
QY 1287 GGAATCGATTACACACCACTCTCTTCTCCACAACAATTCCTACCGCTCTAGCCAT 1346
Db 1443 GGTATTGACACACATCCACTTCAATTCACCCCAACAAATCCCTACAGCTTTTGAAGC 1384
QY 1347 GCGTTTACAAACATGTTTGGAGCGGCACAGGTCCTGTAGATGGCAGCAATGCGACTACT 1406
Db 1383 GGTGTAACAACCATGATTTGGTGGGAACTGGTCTGCTGATGGCACTAATGGCACTACT 1324
QY 1407 ATCACTCCGGGCAAAATGGAACCTTGACCGCATGTTTGGCGGAGCAGAGAGATTTCTATG 1466
Db 1323 ATCACTCCAGGACAGAAATTTAAATGGATGCTCAGAGCGGCTGAAGATATTCTATG 1264
QY 1467 ATGTGGGCTTTTGGGAAAGCAATAGCTCTAGCAAAAACAACTTGTAGAACAACTA 1526
Db 1263 AACTTAGGTTTCTTGGCTAAAGTAACTGCTTTAAACGACGCGAGCTTAGCCGATCAAT 1204
QY 1527 GAAGCGGCGGATTTGTTTAAATTTGATGATGAGCTGGGGCACACACCAAGTCGATC 1586
Db 1203 GAAGCTGGTGATTTGGCTTTAAATCCACGAGACTGGGGCACCACTCTTCTGCAATC 1144
QY 1587 GATCACTGCTGAGCGGTGACAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACA 1646
Db 1143 AATCATGCGTTAGATGTTGCAGCAATACGATGTGCAAGTGCCTATCCACACGACAT 1084
QY 1647 GTCAATGAGCGAGTTATGATGATGACACCCCTAAATGCAATGAACGGCGCGCCATCAT 1706
Db 1083 TTGAATGAAGCGGTTGCGTGAAGACACTATGGCAGCTATTGCCGGACGCACTATGCAC 1024
QY 1707 GCCTACCACTTGAAGGAGCGGTGAGGACACTCACTGATGTTATCACCATGGCAGCG 1766
Db 1023 ACTTTCCACACTGAAGGTGCTGGCGGACACGCTCTCTGATATTAATGAAGTAGCTGT 964
QY 1767 GAGCTCAATATTCTACCTCTCCACCCACCCCACTATTCCCTATACCAATTAATACGGTT 1826
Db 963 GAACACAACTTCTCCCGCTCCACTAACCCCACTATCCCTTCACTGTGAATACAGAA 904
QY 1827 GCAGAACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGCAATCCGCGAGAT 1886
Db 903 GCAGAACACATGACATGCTTATGTTGTGCCACCACTTGGATAAAAGCATTAAGAAGAT 844
QY 1887 TTACAAATTTTCAAGCGGATTCGCGCCCGGCTCTATCGCGGCTGAAGATGTCCTCAT 1946
Db 843 GTTCAGTTGCTGATTTCAAGGATTCGCGCCCTCAAAACCAATTCGCGCTGAAGACACTTTGCAT 784
QY 1947 GATATGGGTGATCGCGATGACAGCTCGGATTCGCAAGCAATGGGGGTGCGAGCGAA 2006
Db 783 GACATGGGATTTTCTCAATCAACAGCTCTGACTCTCAAGCTATGGGTCTGTGGGTGAA 724
QY 2007 GTGATTTCTTCGAATTTGGCAGACTGCGGATAGAAATAAAAAAGAAATTTGGTAAAGCTTCTC 2066
Db 723 GTTATCACTAGAACTTGGCAACAGCTGACAAACAAACAAAGAAATTTGGCCGCTTGAA 664

QY 2067 GAAGATGGCAAGATAAACGATAATTTCCGATTAAGCGCTACATCTCAAATAACATATC 2126
Db 663 GAAGAAAAAGCGGATAAACGAACTTTAGGATCAACGCTACTTGTCTAAATAACCACTT 604
QY 2127 AACCCCGCTTTGACCCACGCGGTGAGGAGTATATCGGCTCTGTGGAGAGGGCAAGATC 2186
Db 603 AACCCAGCGATCGCTCATGGGATTAGCGAGTATGTAGGTTCTGTAGAGTGGGCAAGTG 544
QY 2187 GCGGACTTGGTGTGTGGAAATCTGCTTTTGGCGTAAACCCAAATCGTGATCAAA 2246
Db 543 GCTGACTTGGTATTTGGAGTCCGCAATTTTGGCGTAAACCCCAACATGATCATCAAA 484
QY 2247 GCGGATATGGTGTCTTCTGAAATGGGCGATTTCTAACCGCTGTGTGCCACATCCCAA 2306
Db 483 GCGGGTTCATTCGGTTGAGTCAATGGGTGACGCAACGCTTCTATCCCTACCCCAA 424
QY 2307 CCGGTTTATACCGCAATGTTTGGGCATCAGCGCAGCGGAATTTGACACAGATC 2366
Db 423 CCAAGTTTATACAGAGAAATGTTTCGCTCATCATGGTAAAGCCAAATACGATGCAACATC 364
QY 2367 ACTTTGTTTCCAAAGTCGCTATGAAATGGGTGAAAGAAAGCTGGGCTTAGAGGCG 2426
Db 363 ACTTTTGTGCTCAAGCGCTTATGACAAAGGATTTAAAGAGAAATTAGGCTTTGAAGA 304
QY 2427 CRAAGTTCTACCGGTCAAAAACCTGCCGTAAACATCAACCAAGAAAGACTTTCAAGTTCAACGAC 2486
Db 303 CRAAGTTGCGGTAAAAAATTCGAGAAACATCACTAAAAAAGACATGCAATTTCAACGAC 244
QY 2487 AAAACGGCAAAATCAACGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTC 2546
Db 243 ACTACGCTCACATTTGAAGTCAATCTCTGAAACTTACCATGTGTCGTGGATGGCAAGAA 184
QY 2547 TGCACCTCTAAACCCACCTCGCAAGTCCTCTAGCCAGCGCTACACTTTCTCTAGG 2604
Db 183 GTAACTTCTAAACACGCAATAAAGTGAAGTGGCGCACTCTTTAGCAITTTCTAGG 126

RESULT 15
AR054309

LOCUS AR054309 2735 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5837240.
ACCESSION AR054309
VERSION AR054309.1 GI:5979886
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2735)
AUTHORS Lee, C.K., Monath, T.P., Ackerman, S.K., Thomas, W.D., Soman, G.,
Kleanthous, H., Weltzin, R.A., Pappo, J., Ermak, T., Guirakhoo, P.,
Bhagat, R. and Sussman, I.

TITLE Multimeric, recombinant urease vaccine
JOURNAL Patent: US 5837240-A 1 17-NOV-1998;
FEATURES Location/Qualifiers

source

1..2735
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 37.3%; Score 1075.2; DB 6; Length 2735;
Best Local Similarity 66.0%; Pred. No. 2.4e-204;
Matches 1622; Conservative 0; Mismatches 798; Indels 38; Gaps 3;

QY 182 AAAATTTAAACACAGGAGTAATAGGTGAAACTCACACCCCAAGAGCAAGAAAGTTCTTG 241
Db 78 ACACCTTTAAGATAGGAGATGAGATGAAACTCACCCCAAAAGAGTTAGATTGATG 137
QY 242 TTATATTTATCGCGGCGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTTAAGCTCAACCAA 301
Db 138 CTCACACTCGCTGGAGATTTGGCTAAAAAACGCAAGAAAGAGGCATTAAAGCTTAAC 197
QY 302 CCCGAAGCCATTGCTTTACATTAGTGCCCATATTATGACGAAGCGCGCGGTGAAAAA 361

Qy	2487	AAAAACGGCAAAATCACCCTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTC	2546
Db	2415	ACTACTGCTCACATTGAAGTCAATCCTGAAACTTACCATGTGTTTCGTGGATGGCAAGAA	2474
Qy	2547	TGCACCTCTAAACCCACCTCGCAAGTGCCTCTAGCCCGCGCTACACTTTCTTCTAGG	2604
Db	2475	GTAACCTTCTAAACCGAGCAATAAAGTGAGCTTGGCGCAACTCTTTAGCATTTTCTAGG	2532

Search completed: November 29, 2005, 04:30:37
 Job time : 14066 secs

XX The present invention relates to a novel *Helicobacter felis* urease X and
CC y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against *Helicobacter felis* infections
CC and in diagnostic tests to detect antibodies against *Helicobacter felis*.
CC *Helicobacter felis* is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and y subunits
CC in the manufacture of vaccines. The present sequence represents an urease
CC subunit polypeptide complex encoding sequence of the invention.
XX
SQ Sequence 2883 BP; 871 A; 642 C; 666 G; 698 T; 0 U; 6 Other;

Query Match	99.9%	Score 2880.6	DB 7	Length 2883
Best Local Similarity	100.0%	Prod. No. 0		
Matches 2883	Conservative	0	Mismatches	0
			Indels	Gaps
Qy	1	RGRGAGATTTTCCARCACTTCAAGCACATATTGATCCTGTGTTGGTGGTGTAAATTTCR	60	
Db	1	RGRGAGATTTTCCARCACTTCAAGCACATATTGATCCTGTGTTGGTGGTGTAAATTTCR	60	
Qy	61	ACTTGTAAATCTATTATTAATTTTTTAAATAATTTACTTTATTATCATATATAATAATATTA	120	
Db	61	ACTTGTAAATCTATTATTAATTTTTTAAATAATTTACTTTATTATCATATATAATAATATTA	120	
Qy	121	TTACTTATATTAATAAGTTAATAAAAAAGTAACGAAATTTAGGACTATATAATCCCATTCGCTT	180	
Db	121	TTACTTATATTAATAAGTTAATAAAAAAGTAACGAAATTTAGGACTATATAATCCCATTCGCTT	180	
Qy	181	TAAATTTTACACAGGAGTAATAGGTGAARACTCACCCAAAGAGCAAGAAAGTCTT	240	
Db	181	TAAATTTTACACAGGAGTAATAGGTGAARACTCACCCAAAGAGCAAGAAAGTCTT	240	
Qy	241	GTTTATATTATCGCGGCGAAGTGGCTAGAAAGCGCAAAGCAGAGGCTTTAAAGCTCAACCA	300	
Db	241	GTTTATATTATCGCGGCGAAGTGGCTAGAAAGCGCAAAGCAGAGGCTTTAAAGCTCAACCA	300	
Qy	301	ACCCGAAGCCATTGCTTACATTAGTGCCCATATTATGACGAAGCGCGGTGGGAAAAA	360	
Db	301	ACCCGAAGCCATTGCTTACATTAGTGCCCATATTATGACGAAGCGCGGTGGGAAAAA	360	
Qy	361	AACCGTTGCCAGCTTATGGAAGAGTGCATGCATTTTGTGAAAAAGATGAATGCTT	420	
Db	361	AACCGTTGCCAGCTTATGGAAGAGTGCATGCATTTTGTGAAAAAGATGAATGCTT	420	
Qy	421	CGGGTGGGTAAATGTTTCCGATCTAGGTGTAGAAGCCACTTTTCTGTAGTGTACGAA	480	
Db	421	CGGGTGGGTAAATGTTTCCGATCTAGGTGTAGAAGCCACTTTTCTGTAGTGTACGAA	480	
Qy	481	ACTTGTAACTGTGAATTTGGCCCATCGAAACGAGATGAGCACTTCAAAGCGGCGAAAGTAA	540	
Db	481	ACTTGTAACTGTGAATTTGGCCCATCGAAACGAGATGAGCACTTCAAAGCGGCGAAAGTAA	540	
Qy	541	ATTTGGTTGCGATTAAGACATCGAGCTCAATGCAGGCGAAAGAGTAAACGAACTTGAAGT	600	
Db	541	ATTTGGTTGCGATTAAGACATCGAGCTCAATGCAGGCGAAAGAGTAAACGAACTTGAAGT	600	
Qy	601	TACTTAAATGAAGGCTAAATCTTGCATGTGGTAGCCATTTTCCACTTCTTTTGAAGCTAA	660	
Db	601	TACTTAAATGAAGGCTAAATCTTGCATGTGGTAGCCATTTTCCACTTCTTTTGAAGCTAA	660	
Qy	661	CAAGGCATATAAATTCGATCGTGAAAAAGCCCTATGCGCAACGCTAGATATTTCCCTCTGG	720	
Db	661	CAAGGCATATAAATTCGATCGTGAAAAAGCCCTATGCGCAACGCTAGATATTTCCCTCTGG	720	
Qy	721	CAACACGCTAGCATTTGGGCGAGGACAAACCCGCGAAAGTGCAGTTGATTTCTCTTGGTGG	780	
Db	721	CAACACGCTAGCATTTGGGCGAGGACAAACCCGCGCGAAAGTGCAGTTGATTTCTCTTGGTGG	780	
Qy	781	CAGTAAAAAAGTGATTGCGATCAACCGGCTTGTGAATAACATCGCGGATGAAGCCGATAA	840	
Db	781	CAGTAAAAAAGTGATTGCGATGAAACGGGCTTGTGAATAACATCGCGGATGAAGCCGATAA	840	
Qy	841	ACATAAAGCGCTTGCAAGGCGCAAACTCTCAACGATTTTATCAAGTAGAGGAGATCTCCCATGA	900	

Db 1921 CTATCGCGCTGAAGATGTCTCCATGATATGGGTGTGATCGCGATGACAAGCTCGGATT 1980
Qy 1981 CCAGAGCAATGGGGCTGCAGGCGAAGTGAATTCCTCGAACTTTGGCAGACTGCGGATAAGA 2040
Db 1981 CGCAGCAATGGGGCTGCAGGCGAAGTGAATTCCTCGAACTTTGGCAGACTGCGGATAAGA 2040
Qy 2041 ATAAAAAGAAATTTGGTAAAGTCTCTGAAGATGGCAAGATAACGATAATTTCCGCAATTA 2100
Db 2041 ATAAAAAGAAATTTGGTAAAGTCTCTGAAGATGGCAAGATAACGATAATTTCCGCAATTA 2100
Qy 2101 AGCGGTACATCTCCAAATACATATCAACCCCGCTTTGACCCACGCGGTGAGCGAGTATA 2160
Db 2101 AGCGGTACATCTCCAAATACATATCAACCCCGCTTTGACCCACGCGGTGAGCGAGTATA 2160
Qy 2161 TCGGCTCTGTGAAGAGGCAAGATCGCGCACTTGGTGTGTGGAATCTCGCTTTTGTG 2220
Db 2161 TCGGCTCTGTGAAGAGGCAAGATCGCGCACTTGGTGTGTGGAATCTCGCTTTTGTG 2220
Qy 2221 GCGTAAACCCAAAATCGTGATCAAGGCGGTATGGTGTCTTCTGAAATGGGCGATT 2280
Db 2221 GCGTAAACCCAAAATCGTGATCAAGGCGGTATGGTGTCTTCTGAAATGGGCGATT 2280
Qy 2281 CTAACGCGTCTGTGCCACTCCCAACCGGTTTATTACCGCGAAATGTTTGGGCAATCACG 2340
Db 2281 CTAACGCGTCTGTGCCACTCCCAACCGGTTTATTACCGCGAAATGTTTGGGCAATCACG 2340
Qy 2341 GCAAGCGGAAATTTGACACCGACATCACATTTTGTTCCTCAAGTCGCTATGAAATGGCG 2400
Db 2341 GCAAGCGGAAATTTGACACCGACATCACATTTTGTTCCTCAAGTCGCTATGAAATGGCG 2400
Qy 2401 TGAAGAAAGCTGGGCTTTAGAGCGCAAGTCTTACCGGTCAAAAACCTGCCGTAAATCA 2460
Db 2401 TGAAGAAAGCTGGGCTTTAGAGCGCAAGTCTTACCGGTCAAAAACCTGCCGTAAATCA 2460
Qy 2461 CCAAGAAAGCTTCAAGTTCAACGACAAACCGGCAAAAATCAACCGTCGATCCGAAACCT 2520
Db 2461 CCAAGAAAGCTTCAAGTTCAACGACAAACCGGCAAAAATCAACCGTCGATCCGAAACCT 2520
Qy 2521 TCGAGGTCTTTAGATGGCAACTCTGACCTCTAAACCCACCTCGCAAGTCCTCTAG 2580
Db 2521 TCGAGGTCTTTAGATGGCAACTCTGACCTCTAAACCCACCTCGCAAGTCCTCTAG 2580
Qy 2581 CCCAGCGCTACACTTTCTTAGGCAACAATGCCCTTTGGGGCGCAGGTATTTTAGGAA 2640
Db 2581 CCCAGCGCTACACTTTCTTAGGCAACAATGCCCTTTGGGGCGCAGGTATTTTAGGAA 2640
Qy 2641 TCTTCATCAACGCACTCGATCGGTCTTGGGTGTGCGATCGTGTCTTTAAACCAAC 2700
Db 2641 TCTTCATCAACGCACTCGATCGGTCTTGGGTGTGCGATCGTGTCTTTAAACCAAC 2700
Qy 2701 TTTTCATCTTTAAGCAATCGCATTTTAAATTAATTTAATTTCTTATTAATTAATATATAT 2760
Db 2701 TTTTCATCTTTAAGCAATCGCATTTTAAATTAATTTAATTTCTTATTAATTAATATATAT 2760
Qy 2761 TATGCCCTCAATTTTAAAGAGAAATATGCGTAGTCTTTGGTATTTGCTATTTGGGG 2820
Db 2761 TATGCCCTCAATTTTAAAGAGAAATATGCGTAGTCTTTGGTATTTGCTATTTGGGG 2820
Qy 2821 TTGTTTGGTGTGGCGCAAGGGTATTGAAACCCATCGCTCAAAAAAGTAGAAGCCAC 2880
Db 2821 TTGTTTGGTGTGGCGCAAGGGTATTGAAACCCATCGCTCAAAAAAGTAGAAGCCAC 2880
Qy 2881 AGG 2883
Db 2881 AGG 2883

RESULT 2
ADJ58249
ID ADJ58249 standard; DNA; 2452 BP.
XX
AC ADJ58249;
XX

DT 06-MAY-2004 (first entry)
XX
DE UreaseXY subunit encoding sequence #4.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX
OS Helicobacter felis.
XX
FH Key Location/Qualifiers
CDS 48..728
FT /*tag= a
FT /product= "urease protein"
FT 739..2445
FT /*tag= b
FT /product= "urease protein"
XX
PN EP11716192-A2.
XX
PD 30-JAN-2002.
XX
PP 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
XX WPI; 2002-124384/17.
DR P-PSDB; ADJ58250, ADJ58251.
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
XX vaccines.
XX
PS Disclosure; SEQ ID NO 13; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseXY subunit of the invention.
XX
SQ Sequence 2452 BP; 746 A; 568 C; 590 G; 548 T; 0 U; 0 Other;

Query Match 77.3%; Score 2228; DB 7; Length 2452;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

Qy 159 AGACTATATCCCATTCCTTTAAATTTTAAACAAGGAGTATAGGTGAACTCACAC 218
Db 1 AGACTATATCCCATTCCTTTAAATTTTAAACAAGGAGTATAGGTGAACTCACAC 60

Qy 219 CCAGAGCAAGAAAGTCTTGTATATTATCGGCGAAGTGGCTAGAAAGCGCAAG 278
Db 61 CCAGAGCAAGAAAGTCTTGTATATTATCGGCGAAGTGGCTAGAAAGCGCAAG 120

Qy 279 CAGAGGCTTAAAGCTCAACCAACCGAGCCATTGCTTACATTAGTCCCATATTATGG 338
Db 121 CAGAGGCTTAAAGCTCAACCAACCGAGCCATTGCTTACATTAGTCCCATATTATGG 180

Qy 339 ACGAAGCGCGCTGGGAAACCCGTTCCCGAGTTATGGAGAGTGCATGACATTTT 398
Db 181 ACGAGCGCGCTGTGGCAAAAAACCGTTCCGGAACCTTATGGAAGAGTGTATGCATTTT 240

Qy 399 TGAAGAAAGTGAAGTAATGCCCGGGTGGTAAATGTTCCGATCTAGGTCTAGAAG 458
Db 241 TGAAGAAAGTGAAGTGAATGCCCGGGTGGGAAATATGGTCCCTGATTTGGCGGTGAAG 300

Qy 459 CCACCTTTCTGTGATGATCGAAACTTGTAACTGTGAATTGGCCCATCGAACCATGAGC 518

[illegible]

RESULT 3
ADJ58240

ID ADJ58240 standard; DNA; 2405 BP.
AC ADJ58240;
XX
XX 06-MAY-2004 (first entry)
DT
XX
DE UreaseXY subunit encoding sequence #1.
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX Helicobacter felis.
OS
XX
FH Key Location/Qualifiers
FT CDS 1. .681
FT /*tag= a
FT /product= "urease protein"
FT 692. .2398
FT /*tag= b
FT /product= "urease protein"
XX
XX EP1176192-A2.
PN
XX
PD 30-JAN-2002.
XX
XX 11-JUL-2001; 2001EP-00202666.
PF
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
XX (ALKU) AKZO NOBEL NV.
PA
XX
PI Kusters JG, Cattoli G;
XX
XX WPI: 2002-124384/17.
DR P-PSDB; ADJ58241, ADJ58242.
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
PT vaccines.
XX
PS Disclosure; SEQ ID NO 4; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseXY subunit of the invention.
XX
SQ Sequence 2405 BP; 715 A; 571 C; 588 G; 531 T; 0 U; 0 Other;

Query Match 74.9%; Score 2158.6; DB 7; Length 2405;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 206 GTGAACCTCACACCCAAAGAGCAAGAAAGTTCTTGTATATATATCGCGGCGAAGTGGCT 265
DB 1 GTGAACCTCACACCCAAAGAGCAAGAAAGTTCTTGTATATATATCGCGGCGAAGTGGCT 60
QY 266 AGAAAGCGCAAGCAGAGGCGCTTAAAGCTCAACCAACCCGAGCCGCTTATGGAAGAG 325
DB 61 AGAAAGCGCAAGCAGAGGCGCTTAAAGCTCAACCAACCCGAGCCGCTTATGGAAGAG 120
QY 326 GCCCATATATGGAAGAGCGCGCGCTTGAAAGAAACCGTTGCCGAGCTTATGGAAGAG 385
DB 121 GCCCATATATGGAAGAGCGCGCGCTTGAAAGAAACCGTTGCCGAGCTTATGGAAGAG 180
QY 386 TGCATGCACCTTTTGAAGAAAGATGAAGTAATGCCCGGGGTGGTAAATATGTTCCCGAT 445
DB 181 TGTATGCACCTTTTGAAGAAAGATGAGGTGATGCCCGGTGGGGAATATGGTCCCTGAT 240
QY 446 CTAGGTGTAGAGCCACCTTTCTGTATGTCGAAACTTGTAACTGTGAACTGGCCCATC 505

DB 241 TTGGCGGTAGAGCCACTTTCCCGCATGGCACCACAACTCGTAACCGTGAATGGCCCAT 300
QY 506 GAACCAAGATGAGCACTTCAAGCGGCGAAGTGAATTTGGTTGCGATAAGACATCGAG 565
DB 301 GAACTGTATGAACACACTTTAAAGCGGCGTGAAGTGAATTTGGCTGTGATAAGACATGAG 360
QY 566 CTCAATGTCAGGCAAGAAAGTAACCGAACTTTGAGGTTACTAATGAAGGCGCTTAAATTCCTTG 625
DB 361 CTCAACGCGGTAAGGAAGTTACCGAGCTTTGAAGTTACCAACGAAGGACCTTAAATTCCTTG 420
QY 626 CATGTGGGTAGCCATTTCACCTCTTTTGAAGCTTAAACGAAGCACTTAAATTCGATCGTGAA 685
DB 421 CATGTGGGTAGCCATTTCACCTCTTTTGAAGCAACCAACGAAGCACTTGAATTCGATCGGAA 480
QY 686 AAGAGCTTATGGCAACCGCTTAGATATTCCTCTGGCAACCGCTACCGATTTGGGCGCAGGA 745
DB 481 AAGAGCTTATGGCAACCGCTTAGATATTCCTCTGGCAACCGCTACCGATTTGGGCGCAGGA 540
QY 746 CAAACCCGCAAGTGCAGTTGATTTCCTCTTTGGTGGCAGTAAACAAAGTGAATTTGGCATGAAC 805
DB 541 CAAACCCGTAAGTGCAGTTAATCCCTCTTTGGCGGTAGTAAACAAAGTGAATTTGGCATGAAC 600
QY 806 GGGCTTTGTAATAACATCGCGGATGAACGCCATAAACATAAAGCGCTTTGCAAGCGGCAAA 865
DB 601 GGGCTTTGTAATAATAATTCGGGACGAACGCGCATAAACACAAAGCACTAGACAAGGCAAAA 660
QY 866 TCTCAGGATTTATCAAGTAAGGAGACTCCCATGAAATGAAACAAACAAAGTATGTAAA 925
DB 661 TCTCAGGATTTATCAAGTAAGGAGACTCCCATGAAATGAAACAAACAAAGTATGTAAA 720
QY 926 TACTACGGACCCACCAAGGCGATAAAGTCGCTTAGGAGATACCGCATCTTTGGGCGAGA 985
DB 721 CACCTACGAGCCACCAAGCGCATTAAGTTCGCTTAGGAGATACCGCATCTTTGGGCGAGA 780
QY 986 AGTAGAACATGACTATACCACTTATGCGCAAGCACTTTAAATTTGGCCGCGGTAAAACTAT 1045
DB 781 AGTAGAACATGACTATACCACTTATGCGCAAGAGCTCAAATTTGGCGCGGTAAAACTAT 840
QY 1046 CGGTAGGGTATGGGTGAGAGCAATAGCCCTGTATGATAACCAACCCCTAGATTAGTCATCAC 1105
DB 841 CCGTGAGGGTATGGGTGAGAGCAATAGCCCTGTATGATAACCAACCCCTAGATTAGTCATCAC 900
QY 1106 TAACGCGATGATTATGACTACACCGGATTTTCAAGCCGACATTCGGATTTAAAAACGG 1165
DB 901 CAAAGCGATGATTATGACTACACCGGATTTTAAAGCCGACATTTGTTTAAAAATGG 960
QY 1166 CAAAATCCATGCGATTGGCAAGGCGAGGAAACAAAGGACATGCAAGATGCGGTAAAGCCCTCA 1225
DB 961 CAAAATCCATGCGATTGGCAAGGCGGGAACAAAGACATGCAAGATGCGGTAAAGCCCTCA 1020
QY 1226 TATGTGCTGGGTGAGGCGACAGAACCATAGCAGGGGAAAGGTATGATTAATACCGCTGG 1285
DB 1021 TATGTGCTGGGTGAGGCGACAGAACCATAGCAGGGGAAAGGTATGATTAATACCGCTGG 1080
QY 1286 GGGATTCGATTACACACCCACTTCTCTTCTCCACAACTTCCCTACCGCTCTAGCCAA 1345
DB 1081 GGGATTCGATTACACACCCACTTCTCTTCTCCACAACTTCCCTACCGCTCTAGCCAA 1140
QY 1346 TGGCGGTACAAACCATGTTTGGAGGCGGACAGGTCTCTGTAGATGGCAAGTGGCACTAC 1405
DB 1141 TGGGTGTACAAACCATGTTTGGAGGTGGCACAGGTCCGGTAGATGGCAAGTGGCACTAC 1200
QY 1406 TATCACTCCGGGCAAAATGGAACCTTGACCGCATTTGTCGCGCAGCAGAGAGTATTTCTAT 1465
DB 1201 CATCACTCCGGGCAAAATGGAACCTTGACCGCATTTGTCGCGCAGCTGAGAGTATTTCTAT 1260
QY 1466 GAATGTGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAACAACTTTCTAGAACAGT 1525
DB 1261 GAATGTGGCTTTTGGGCAAAAGGCAATAGCTCTAGCTAAAAACAACTTCGTAGAACAGT 1320
QY 1526 AGAAGCGGCGCGATTGGTTTTTAAATTTGATGAAGACTGGGGGCAACACCAAGTGGCAT 1585

Db 1321 AGAAGCGGCGGCGATTGGCTTTTAAATTGCGATGAAGACTGGGGCACAACACCAAGTCGGAT 1380
QY CGATCACTGCTTGAGCGTGGCAGATGAATACAGATGTCGAAGTTTGTATCCACACCGATAC 1645
Db 1381 CGATCACTGCTTGAGCGTAGCAGATGAATACAGATGTCGAAGTTTGTATCCACACCGATAC 1440
QY 1646 AGTCAATAGGCGAGGTTATGTAGATGACACCCCTAAATGCAATGAACCGGGCGGCCATCCA 1705
Db 1441 GGTCAATAGGCGAGGTTATGTAGATGACACCCCTAAATGCGATGAACGGGCGGCCATCCA 1500
QY 1706 TGCCTACCACTTAGGCGGCGGTGGAGGACACTCACCTGATGTTATCACCATGCGAGG 1765
Db 1501 TGCCTACCACTTAGGCGGCGGCGGAGGACACTCACCTGATGTTATCACCATGCGAGG 1560
QY 1766 CGAGCTCAATATTCTACCTCTCCACCAACCCCACTATTCCCTATACCAATTAATACGGT 1825
Db 1561 CGAGCTCAATATTCTACCTCTCCACCAACCCCACTATTCCCTATACCAATTAATACGGT 1620
QY 1826 TGCAGAACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGCAATCCGCGAGGA 1885
Db 1621 TGCAGAACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGCAATCCGCGAGGA 1680
QY 1886 TTTACAAATTTCTCAAGCGGTATCGGCCCGGCTCTATCGCGCTGAAGATGTGCTCCA 1945
Db 1681 TCTCCAGTTTTCCTCAAGCGGTATCGGCCCGGCTCTATCGCGCTGAAGATGTGCTCCA 1740
QY 1946 TGATATGGGTGTGATCGCGATGACAGCTCGGATTCGCAAGCAATGGGGCGTGCAGCGCA 2005
Db 1741 TGATATGGGTGTGATCGCGATGACAGCTCGGATTCGCAAGCAATGGGGCGGCTGGGGA 1800
QY 2006 AGTGATTCCTCGAATTTGGCAGACTCGGATAGAAATTAAGAAATTTGGTAAAGCTTCC 2065
Db 1801 AGTGATTCCTCGAATTTGGCAGACTCGCAGACAGAAATTAAGAAATTTGGTAAAGCTTCC 1860
QY 2066 TGAAGATGGCAAGATTAACATTAATTTCCGCATTAAGCGCTACATCTCCAAATACACTAT 2125
Db 1861 TGAAGATGGTGCAGATTAATGACAACTTCGCGATCAAAACGCTATATCTCCAAATACACCAT 1920
QY 2126 CAACCCCGCTTTCACCCACCGCGTGAAGGATATATCGGCTCTGTGGAAGAGGCGAAGAT 2185
Db 1921 TAATCCCGCTTTCACCCACCGCGTGAAGGATATATCGGCTCTGTGGAAGAGGCGAAGAT 1980
QY 2186 CGCGGACTTGGTGTGTGAATCTGCTCTTTTGGGTAAAAACCCCAAAATCGTGATCAA 2245
Db 1981 CGCGGACTTGGTGTGTGAATCTGCTCTTTTGGGTAAAAACCCCAAAATCGTGATCAA 2040
QY 2246 AGCGGTATGGTGTCTTCTGAAATGGGCGATTTAAACGCGTCTGTGCCCACTCCCA 2305
Db 2041 AGCGGTATGGTGTGTCTTCTGAAATGGGCGATTTCTAAACGCGTCTGTGCCCACTCCA 2100
QY 2306 ACCGGTTTATACCGGAAATGTTGGGCGATCAAGGCAAGCGAAATTTGACACCGACAT 2365
Db 2101 GCGGGTTTATACCGGAAATGTTGGGCGATCAAGGCAAGCGAAATTTGACACCGACAT 2160
QY 2366 CACTTTTGTTCCTCAAGTCGCTATGAAATGGCGTGAAGAAAGAAAGCTGGGCTTAGAGCG 2425
Db 2161 CACTTTTGTTCCTCAAGTCGCTATGAAATGGCGTGAAGAAAGAAAGCTGGGCTTAGAGCG 2220
QY 2426 CCAAGTTTCTACCGGTCAAAAATGCGCGTAAACATCAACCAAGAAAGACTTCAAGTTCAACGA 2485
Db 2221 CAAGGTGTACCGGTGAAGAAATGCGCGCAACATCACTAAGAAAGACTTCAAAATCAACAA 2280
QY 2486 CAACACGGCAAAATCAACCGTGCATCCGAAACCTTCGAGGTCTTTGTAGATGCGCAACT 2545
Db 2281 CAACACGGCGCATATCATCTGCGATCTCTAAACCTTCGAGGTCTTTGTAGATGCGCAACT 2340
QY 2546 CTGCACCTCTAAACCCACCTCGCAAGTGCCTCTAGCCAGCGCTACACTTTCTTCTAGGC 2605
Db 2341 CTGCACCTCTAAACCCCGCTCTGAAGTGCCTCTAGCCCAACGCTACACTTTCTTCTAGGC 2400
QY 2606 ACAAT 2610
Db 2401 ACAAT 2405

RESULT 4

ADJ58246
ID ADJ58246 standard; DNA; 2407 BP.
XX
AC ADJ58246;
XX
DT 06-MAY-2004 (first entry)
XX
DE UreaseXY subunit encoding sequence #3.
XX immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX
OS Helicobacter felis.
XX
FH Location/Qualifiers
CDS 2..682
FT /*tag= a
FT /product= "urease protein"
FT 693..2399
FT /*tag= b
FT /product= "urease protein"
XX
XX EP1176192-A2.
XX
XX 30-JAN-2002.
XX
XX 11-JUL-2001; 2001EP-00202666.
XX
XX 17-JUL-2000; 2000EP-00202565.
XX
XX (ALKU) AKZO NOBEL NV.
XX
XX Kusters JG, Cattoli G;
XX WPI; 2002-124384/17.
XX P-PSDB; ADJ58247, ADJ58248.
XX
XX Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
XX the diagnosis of Helicobacter felis infections and in the preparation of
XX vaccines.
XX
XX Disclosure; SEQ ID NO 10; 76pp; English.
XX
XX The present invention relates to a novel Helicobacter felis urease X and
XX Y subunit polypeptides and immunogenic fragments. The polypeptides are
XX used in the manufacture of vaccines against Helicobacter felis infections
XX and in diagnostic tests to detect antibodies against Helicobacter felis.
XX Helicobacter felis is difficult to grow so it is more convenient to use
XX the expression products of the genes encoding the urease X and Y subunits
XX in the manufacture of vaccines. The present sequence represents an
XX ureaseXY subunit of the invention.
XX
SQ Sequence 2407 BP; 720 A; 574 C; 584 G; 528 T; 0 U; 1 Other;
Query Match 74.9%; Score 2158.6; DB 7; Length 2407;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 155; Indels 0; Gaps 0;
QY 206 GTGAAACTCACACCCCAAGAGCAAGAAAGTTCTTGTATATATGCGGGCGAAGTGGCT 265
Db 2 GTGAAACTCACACCCCAAGAGCAAGAAAGTTCTTGTATATATGCGGGCGAAGTGGCT 61
QY 266 AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGAAAGCATTTGCTTACATTAGT 325
Db 62 AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGAAAGCATTTGCTTACATTAGT 121
QY 326 GCCCATATTATGGACGAGCGCGCGTGGAAAAAACCCTTGGCCAGCTTATGCAAGAG 385
Db 122 GCCCATATTATGGACGAGCGCGCGTGGAAAAAAGACCCGTTCGGAACTTATGGAAGAG 181
QY 386 TGCATGCACCTTTTGTAAAAAAGATGAAGTAATGCCCGGGTGGGTAATATGTTTCCCGAT 445

Db 1203 CATCACTCGCGCAAAATGGAACTTGACCCGCGATGTTGCGCGCAGCTGAAGAGTATTCTAT 1262
QY 1466 GAATGGGGCTTTTGGGCAAAAGGCAATAGCTTAGCAAAAAACAATTTGTAGAACAAAGT 1525
Db 1263 GAATGTAGGCTTTTGGGCAAAAGGCAATAGTTCTAGCAAAAAACAATTTGTAGAACAAAGT 1322
QY 1526 AGAAGCGGGCGGATTTGGTTTAAATTCATGAAGACTGCGGGCAACAACCAAGTCCGAT 1585
Db 1323 AGAAGCGGGCGGATTTGGCTTTAAATTCATGAAGACTGCGGGCAACAACCAAGTCCGAT 1382
QY 1586 CGATCACTGCTTGAAGCTGCGGAGATGAATACGATGTCGAAGTTTGTATCCACCCGATAC 1645
Db 1383 CGATCACTGCTTGAAGCTGCGGAGATGAATACGATGTCGAAGTTTGTATCCACCCGATAC 1442
QY 1646 AGTCAATGAGCGAGGTTATGTAGATGACACACCTTAAATGCAATGAACGGCGGCCATCCA 1705
Db 1443 GGTCAATGAGCGAGGTTATGTAGATGACACACCTTAAATGCAATGAACGGCGGCCATCCA 1502
QY 1706 TGCCTACCACTGAGGGAGCGGGTGGAGGACACTCACCTGATGTTTATCACCATGGCAGG 1765
Db 1503 TGCCTACCACTGAGGGAGCGGGCGGAGGACACTCACCTGATGTTTATCACCATGGCAGG 1562
QY 1766 CGAGCTCAATATTTACCTCTCCACACCCCACTATTCCTATACCATTAATACGGT 1825
Db 1563 CGAGCTCAATATTTACCTCTCCACACCCCACTATTCCTATACCATTAATACGGT 1622
QY 1826 TGCAGAACCTTAGACATGCTCATGATGCGACACCTAGACAAACGATCGCGAGGA 1885
Db 1623 TGCAGAACCTTAGACATGCTCATGATGCGACACCTAGACAAACGATCGCGAGGA 1682
QY 1886 TTTACAATTTTCTCAAAGCGTATCGCGCCCGCTCTATCGCGGCTGGAAGATGTGTCTCA 1945
Db 1683 TTTACAATTTTCTCAAAGCGTATCGCGCCCGCTCTATCGCGGCTGGAAGATGTGTCTCA 1742
QY 1946 TGATATGGGTGTATCGGATGACAGCTCGGATTCGGAAGCAATGGGGCGGTGACGGCA 2005
Db 1743 TGATATGGGTGTATCGGATGACAGCTCGGATTCGGAAGCAATGGGGCGGTGACGGCA 1802
QY 2006 AGTGATTCCTCAACTTGGCAGCTGCGGATGAAGATGAAGATTAAGATTTGGTAAAGCTCC 2065
Db 1803 AGTGATTCCTCAACTTGGCAGCTGCGGATGAAGATGAAGATTAAGATTTGGTAAAGCTCC 1862
QY 2066 TGAAGATGCAAGATGAAGATTTCCGCAATTAAGCGCTACATCTCCAAATACACTAT 2125
Db 1863 TGAAGATGTCAGATGAAGATTTCCGCAATTAAGCGCTACATCTCCAAATACACTAT 1922
QY 2126 CAACCCCGCTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGAT 2185
Db 1923 TAATCCCGCTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGAT 1982
QY 2186 CGCCGACTTGGTGTGGAATCCTGCTTTTGGCGTAAACCCCAAAATCGTATCAA 2245
Db 1983 CGCCGACTTGGTGTGGAATCCTGCTTTTGGCGTAAACCCCAAAATCGTATCAA 2042
QY 2246 AGCGGTATGGTGGTCTCTCTGAATGGCGGATTTAAACGGCTGTGTCGCCACTCCCA 2305
Db 2043 AGTGGCATGGTGGTCTCTCTGAATGGCGGATTTAAACGGCTGTGTCGCCACTCCCA 2102
QY 2306 ACCGGTTTATACCGGAAATGTTTGGGCATACCGCAAGGGCGAAATTTGACACCAAGCAT 2365
Db 2103 GCGGGTTTATACCGGAAATGTTTGGGCATACCGCAAGGGCGAAATTTGACACCAAGCAT 2162
QY 2366 CACTTTTGTTCCTCAAGTCG 2385
Db 2163 CACTTTTGTGTCTCAAGCG 2182

RESULT 6
ADQ37847
ID ADQ37847 standard; DNA; 8407 BP.
XX
AC
AC
XX

DT 07-OCT-2004 (first entry)
XX H. bizzozeronii urease gene cluster, ureABIEFGH.
XX
XX Urease; urease gene cluster; urease structural gene;
KW urease accessory gene; ureABIEFGH; Helicobacter bizzozeronii infection;
KW antibacterial; gene; ds.
XX
OS Helicobacter bizzozeronii.
XX
PN US2004142343-A1.
XX
XX 22-JUL-2004.
PD
XX 12-AUG-2003; 2003US-00639273.
XX
XX 16-AUG-2002; 2002US-0404337P.
PR
XX (CHAN/) CHANG Y.
PA (SIMP/) SIMPSON K W.
PA (ZHUJ/) ZHU J.
XX
PI Chang Y, Simpson KW, Zhu J;
XX
XX WPI; 2004-533502/51.
DR GENBANK; AF330621.
XX
XX Novel isolated nucleic acid molecule having urease gene cluster, and
PT conferring on Helicobacter bizzozeronii ability to produce urease, useful
PT as vaccine for preventing disease in mammals infected by H.bizzozeronii.
XX
XX Claim 2; SEQ ID NO 1; 40pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule conferring on
CC Helicobacter bizzozeronii an ability to produce urease, where the nucleic
CC acid molecule is a urease gene cluster comprising at least one urease
CC structural gene and at least one urease accessory gene. The nucleic acid
CC molecule is chosen from ureA, ureB, ureE, ureF, ureG, ureH and ureI. The
CC invention also relates to an isolated protein encoded by the nucleic
CC acid, a vaccine for preventing onset of disease in mammals infected by H.
CC bizzozeronii comprising a nucleic acid and a carrier, and an isolated
CC antibody or its binding portion raised against the nucleic acid. The
CC nucleic acids, proteins and antibodies are useful for vaccinating mammals
CC against onset of disease caused by infection of H. bizzozeronii, which
CC involves administering the sequences. The sequences are useful for
CC detecting H. bizzozeronii in a sample of tissue or body fluids which
CC involves providing a nucleic acid as an antigen, providing an antibody,
CC or providing a nucleotide sequence as a probe in a nucleic acid
CC hybridisation assay, contacting the sample with the antigen or the probe,
CC and detecting any reaction which indicates that H. bizzozeronii is
CC present in the sample. This sequence represents the H. bizzozeronii
CC urease gene cluster, ureABIEFGH.
XX
SQ Sequence 8407 BP; 2264 A; 1937 C; 1969 G; 2237 T; 0 U; 0 Other;

Query Match 41.1%; Score 1184.4; DB 12; Length 8407;
Best Local Similarity 68.5%; Pred. No. 1.5e-255;
Matches 1681; Conservative 0; Mismatches 741; Indels 32; Gaps 2;
QY 183 AAATTAAACACAGAGTAGTAATAGTGAAACTCACCCCAAGAGCAAGAAAGTCTTGT 242
Db 2319 AAATTGGTAGAGGAGTTTAGATGAATTAACCCCTAAAGAGCTGACAGCTCATGT 2378
QY 243 TATATTATCGGGCGAAGTGGCTAGAAAGCGCAAGAGGGCTTAAAGTCAACCAAC 302
Db 2379 TGCATTTATCGGGCGAATTTGGCTTAAAAAACGCAAGCAAAATGGCTTAAATATA 2438
QY 303 CGAAGCCATTCCTTACATTAGTCCCATATTATGACGAGCGCGCGTGGAAAAAAA 362
Db 2439 CTGAGGCAGTAGCCCTCATCAGTGCCTCATGTGATGGAAGAGCCCGTGCAGGTAAGAAA 2498
QY 363 CGGTTGCCCGAGCTTATGGAAGAGTGCATGCACATTTTGTAAAAAAGATGAAGTATGCCCG 422

Db	2499	GTGTGGCGGATTGATGCAAGAGCGGACGACACTTCTTAAAGCTGATGATGTCATGCCCG	2558
Qy	423	CGGTGGTAAATATGGTTCCTCGATCTAGGTGTAGAACCCACTTTCCTGATGGTACGAAC	482
Db	2559	GTGTAGCCCATATGATCCACGAAGTGGGATTTGAAGCTAACTTCCCTGATGGGACAAAC	2618
Qy	483	TTGTAACTGTGAATTTGGCCCATCGAACACAGATGAGCACTTCAAAGCGGGGGAAGTGAAT	542
Db	2619	TGGTAAACCATCCATACCCCGTGTGAAGATGGTGGGCATAAATTTGGCTCCGGGTGAAGTGA	2678
Qy	543	TTGGTTGGCCATAAAGACATCGAGCTCAATGTCAGGCAAGAGTAAACCGAACTTGAGGTGA	602
Db	2679	TTTTGAAAACGAGACATCACTTTGAAATGACGAGCAACAGCCACCACTTTAGAAAGTGC	2738
Qy	603	CTAATGAAGGGCCCTAAATCTTGCATGTGGGTAGCCATTTCACATCTTTTGAAGCTAAACA	662
Db	2739	ATAACAAAGCGATCGCCCGTGCAAGTGGCTCCCACTTCCACTTCTTTTGAAGTGAATA	2798
Qy	663	AGGCATAAAATTCGATCGTGAAGGAGCTATGGGCAACGCCCTAGATATTCCTCTGGCA	722
Db	2799	AGCTTTTGGAAATTTGATCGTGAAGGAGCCCTATGGCAAGCCCTAGACATTTGCTTCTGGAA	2858
Qy	723	ACAGCTACGATTTGGGCGAGGACAAACCGCAAGTGCAGTTGATTTCTCTTGGTGGCA	782
Db	2859	CGCTGTGGCTTTTGAACCGGTGAGAAAAAACCCTGGGAATTTGATTCAAATTTGGCGGTA	2918
Qy	783	GTAAAAAGTGATTTGCATGAACGGGCTTGTGAATAACATCGCGGATGAACGCCATAAAC	842
Db	2919	ACCAACGATTTTACGCTTTTAACTCTCTTGTGGATCGCCAAAGCCGATACTGATGGCAAA	2978
Qy	843	ATAAGCGCTTGACAAGGGGAAATCTCACGGATTT-----	877
Db	2979	AACTTGCTCTCAACCGCGCAAGAACATGGCTTTTGGTGTGTGAATTTGGCGTTGGGATA	3038
Qy	878	----ATCAAGTAAGAGACTCCCATGAAATGAAAA--ACAAGAAATGTAATACCT	930
Db	3039	AAAAATAAGAAAGGACATCCCATGAAAAAATCTCTCGAAAGAAATATGTTTCTATGT	3098
Qy	931	ACGGACCCCAAGGCGATAAAGTCGCTTAGAGATACCGATCTTTTGGCGAGAGTAG	990
Db	3099	ATGGACCCACTAGCGGGATAAGTGTAGATTGGGCGATACCGACTGATCTTAGAAGTCG	3158
Qy	991	AACATGACTATACCACTATGGCGAGAACTTAAATTTGGCGGGTAAACTATCCGTG	1050
Db	3159	AACATGACTACCACTTATGGCGAGAAATTAAGTTTGGTGGCGTAAACCACTTCGCG	3218
Qy	1051	AGGGTATGGTACAGCAATAGCCCTGATGAACACCCCTAGATTAGTTCATCACTAACG	1110
Db	3219	ATGGGATGGCAAAACCAACAGCCCCAGACCCAGCACTCGATCTTTGTGCTCACTAACG	3278
Qy	1111	CGATGATATCGACTACACCGGATTTTACAAAGCCGACATTTGGGATTTAAAAACGGCAAA	1170
Db	3279	CCCTGATCGTGGATTACACCGCATTTTAAAGCCGATATTTGGCATTTAAANATGGCAAAA	3338
Qy	1171	TCCATGGCATTTGGCAAGCAGGCAATAAGACATGCAAGATGGCGTTTGGCAACAATCTTT	3398
Db	1231	TCGTGGGTGGGACAGAAAGCACTAGCAGGGGAGGTATGATTTATACCGCTGGGGGAA	1290
Qy	3399	CGGTGGGCCCTGCTACTGAGGCTTTTGGCCGTGAAGGGCTGATTTTACAGCTGGTGGGA	3458
Db	1291	TCCGATTCACACACCCACTTCTTCTCCACAACTTCCCTACCGCTCTAGCCCAATGGCG	1350
Qy	3459	TTGACACCCACATCCACTTTATTTCTCCCAACAAATCCCAACAGCATTTTGGCAGCGGA	3518
Db	1351	TTACAAACCATGTTTGGAGGGGACAGGTCTGTAGATGGCAGCAAGTTCGCACTACTATCA	1410
Qy	3519	TCACAACCATGATTTGGTGGGGACAGGTCCAGCTGATGGACTAACCGACTACCATCA	3578
Db	1411	CTCGGGCAATGGAACCTTGACCGCATTTTGGCGGAGGAGAGATTTCTATGATG	1470
Qy	3579	CTCGGGCGCTGGAACTTTAAAAACCATGCTCCGTGCTCTGGAAGAAATATGCCATGAACT	3638
Db	1471	TGGGCTTTTGGGCAAAAGGCAATAGCTCTTAGCAAAAAAACAATTGTGTAGAACAAAGTAGAAG	1530
Qy	3639	TGGGCTAATTTGGGTAAGGGAATGTGTCTTATGAACCTCCCTGGTGGTCAACTCGAAG	3698
Db	1531	CGGGCGGATTTGGTTTAAATTCGACTGGGCAACAACCAAGTGGCATCGATC	1590
Qy	3699	CTGAGCCCATTTGGCTTTAAATCCAGAACCTGGGTAGCACACCTTGAGCCATCTACC	3758
Db	1591	ACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCA	1650
Qy	3759	ATTGCTTGAATGTGGCTTGACAAATACGATGTGCAAGTGGCTATCCACACCGATACCTTGA	3818
Db	1651	ATGAGCAGGTTATGTATGATGACACCCCTAAATGCAATGAACGGCGGCCCATTCATGCT	1710
Qy	3819	ATGAAGCGGCTGTGTGGAAGACACTTTTGAAGCCATTGTCTGGGCGCACTATCCACACTT	3878
Db	1711	ACCACATTTGAGGAGCGGTGGAGGACACTCACCTGATGTTTATCACCATGCGAGCGAGC	1770
Qy	3879	TCCACACTGAAGTGTGTGGGCGGACCGCTCCGATGTCATTAAGATGCTGGCGAAT	3938
Db	1771	TCATATTTTACCTCTCTCCACACCCCACTATTTCCCTATACCATTAATACGTTGCGAG	1830
Qy	3939	TTAACATCTCCAGCTTTTACCAACCCCACTTCTTTCACCGTGAATACAGAAGCG	3998
Db	1831	AACATTTAGACATGCTCATGACATGCCACCACTTAGACAAAGCATCCGCGAGGATTTAC	1890
Qy	3999	AACACATGGACATGTTGATGTTGCCACCACTTTGGATATAAAACATCAAGAAGATGTCC	4058
Db	1891	AAATTTCTCAAGCGGTATCGCCCGCTCTATCGCGGCTGAAGATGTGCTCCATGATA	1950
Qy	4059	AGTTTGTGATTTCTAGGATTTGGCCCCCAACCATTCGCCGCTGAGGACAAATCTCCACGATA	4118
Db	1951	TGGGTGTGATCGCGATGACAAAGCTTCGGATTCGCAAGCAATGGGGCGTGCAGCGAAGTGA	2010
Qy	4119	TGGGGATTTTCTCTATCACCACTCTGACTCCCAAGCGATGGCGCGTGTAGCGAGGTCA	4178
Db	2011	TTCTCTGAACTTGGCAGACTGGCGATAAGATAAAAAAGAAATTTTGGTAAGCTTCTCTAAG	2070
Qy	4179	TCACCCGCACTTGGCAAAACAGCGACAAAAAACAAGAAATTTTGGTTCGCTCGAGG	4238
Db	2071	ATGGCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTATCAACC	2130
Qy	4239	AAAAAGCGATATATGACACTTCGCACTCAAGCTACATTTCCAAATACACCATCAACC	4298
Db	2131	CCGCTTTGACCCACCGGCTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGTCGCG	2190
Qy	4299	CCGCTATGACACCGCATTTCTGAAATATGTCGGCTCTGTAGAAAGTGGGCAAAATTCGCG	4358
Db	2191	ACTTGGTGGTGGGAATCTGCTCTTTTGGCTTAAACCCCAAAATCTGTGATCAAAAGCG	2250
Qy	4359	ATTTGGTGTCTTGGAGTCTGCTTCTTTGGCATTTAAACCCCAACATGATCATCAAGGCG	4418
Db	2251	GTATGGTGTCTCTCTGAAATGGGCGATTTCTAACGGCTCTGTGCCCATCCCAACCGG	2310
Qy	4419	GATTCATCGCACTTCTCTCAATGGGCGATGCAATTCCTCTATCCCACTCCCAACCGG	4478
Db	2311	TTTATTAACCGGAATGTTTGGGCGATCAGCGCAAGGGGAAATTTTGACACCCAGCATCACTT	2370
Qy	4479	TGATTAACCGGAATGTTTGGCCCACTGTTGAAGCCAAATTTTGACACCAATATCACTT	4538
Db	2371	TTGTTTCCAAAGTCCCTATGAAATGGCGGTGAAGAAAGCGTGGCTTTAGAGCCCAAG	2430
Qy	4539	TTGTTATCCCAAGTGGCTTATGACAAAGGAGTTTAAAGAGAGTTGGGCTTGCAAGAGTGG	4598
Db	2431	TTCTACCGGTCAAAAACTGGCGTAACATCAACAAGAAAGACTTCAAGTTCAACGACAAA	2490
Qy	4599	TTTTGCCAGTTAAAACTGGCGCAACATCAACAAAAAGACCTCAAAATTCACGATGTTA	4658
Db	2491	CGCAAAATCACCGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCA	2550
Qy	4659	CCGACACATCGAAGTCAATCTCTGAAACCTTAAAGTTTAAAGTGGATGGCAAGAGTTA	4718

QY 2551 CCTCTAAACCCACTCGCAGTGCCTCTAGCCAGCGCTACACTTCTCTCTAGG 2604
|||||
Db 4719 CTTCCAAAGCAGCGGATAAAATCAGCTAGCACAACCTCTACAACTTGTCTTAGG 4772
|||||

RESULT 7

ID AAT44351 standard; cDNA; 2735 BP.
XX AAT44351;

XX AC AAT44351;
DT 16-OCT-2003 (revised)
DT 11-FEB-1997 (first entry)
DE H. pylori ureA + ureB gene locus in pORV214.
XX

KW Urease; ureA gene; ureB gene; vaccine; ds.
XX

XX Helicobacter; strain CPM630.
OS unidentified bacteriophage; T7.
OS Chimeric.

PH Key Location/Qualifiers
FT promoter
1. .16

FT /tag= a
FT /note= "T7 promoter provides transcription initiation for
FT the urease genes"
FT 33. .43
FT misc_signal

FT /tag= b
FT /function= "operator"
FT /note= "Lac operator provides inducible expression of the
FT urease genes"
FT complement(46. .67)

FT primer_bind
FT /tag= c
FT /note= "BL1 primer"
FT 102. .818

FT CDS
FT /tag= d
FT /product= "urease A subunit"
FT 822. .2531

FT CDS
FT /tag= e
FT /product= "urease B subunit"
FT 2546. .2569

FT primer_bind
FT /tag= f
FT /note= "BL2 primer"
FT 2693. .2735

FT terminator
FT /tag= g
FT /note= "T7 terminator"
FT

XX WO9633732-A1.

PN 31-OCT-1996.

PD 25-APR-1996; 96WO-US005800.

XX 28-APR-1995; 95US-00431041.

PR 06-DEC-1995; 95US-00568122.

XX (ORAV-) ORAVAX INC.

PA Lee CK, Monath TP, Ackerman SK, Thomas WD, Soman G, Kleanthous H;
PI Weltzin RA, Pappo J, Ermak T, Guirakhoo F, Bhagat H, Sussman I;
DR WPI; 1996-497373/49.

DR P-PSDB; AAW07193, AAW07194.
XX

XX Vaccine for inducing mucosal response to Helicobacter - contg. multimeric
PT urease complex and pref. an antibiotic, anti-secretory agent or bismuth
PT salt.

XX Disclosure; Page 68-70; 98pp; English.

PS A cDNA clone (AAT44351), derived from pORV214, includes the ureA and ureB
XX sequences coding for the urease A (AAW07193) and urease B (AAW07194)
CC

CC subunits of Helicobacter pylori clinical isolate CPM630. To obtain
CC pORV214, a genomic DNA library of CPM630 was screened with anti-
CC Helicobacter urease antibody. A 17 kb SalI fragment from an isolated
CC clone was subcloned into pUC18 to give pSCP1. PCR primers (AAT44352-53)
CC were used to amplify a 2.5 kb fragment from pSCP1, which was inserted
CC into pET24+ to give pORV214. The vector was utilised in the prodn. of
CC recombinant, enzymatically inactive, multimeric urease in E. coli
CC transformants for use in vaccines to treat or prevent Helicobacter
CC infection. (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 2735 BP; 875 A; 562 C; 635 G; 663 T; 0 U; 0 Other;

Query Match 37.3%; Score 1075.2; DB 2; Length 2735;
Best Local Similarity 66.0%; Pred. No. 3.Se-231;
Matches 1622; Conservative 0; Mismatches 798; Indels 38; Gaps 3;

QY 182 AAATTTACACAGGAGTAATAGTGAACCTCACACCCCAAGAGCAGAAAGCTCTTCTTG 241
|||||
Db 78 ACACCTTTAAGATAGGAGATGAGATGAACCTCACCCCAAGAGTGTAGTAAGTTGATG 137
|||||
QY 242 TTATATTATGCGGCGAAGTGGCTAGAAAGCGAAAGCAGAGGCTTTAAAGCTCAACCAA 301
|||||
Db 138 CTCCTACTGCTGAGATTTGGCTTAAACACGCAAGAAAGGCACTTAAGCTTAAGTAT 197
|||||
QY 302 CCCGAAGCCATTGCTTACATTAGTGCCCATATTATGACGAAAGCGCCGCTGGAAAAAAA 361
|||||
Db 198 GTAGAAGCAGTAGCTTTGATTAGTGCCCATATTATGGAAGAGAGAGCTGGTAAAAAG 257
|||||
QY 362 ACCGTTGCCAGCTTATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGTATGCC 421
|||||
Db 258 ACTGCGGCTGAATTGATGCAAGAGGCGCAGCTCTTTTAAACACGATGATGTGATGAT 317
|||||
QY 422 GGGGTGGTAATATGTTCCCGATCTAGGTGTAGAGCCACCTTTCTCTGATGGTACGAAA 481
|||||
Db 318 GCGGTGGCAAGCATGATCCATGAAGTGGGTATTGAACGATGTTTCTCTGATGGACTAAA 377
|||||
QY 482 CTTGTAACTGTGAATTTGGCCCATCGAACCAAGATGAGCACTTCAAAGCGGCGAAGTGA 541
|||||
Db 378 CTCGTAACCGTGCATACCCCTATTGAGGCCAATGGTAAATTAGTTCCTGGTGGTTG-- 434
|||||
QY 542 TTTGGTTGGATAAAGACATCGAGCTCAATGCGAGGCAAGAAAGTAAACCGAACTTGA 601
|||||
Db 435 TTCTTAAAAAATGAAGACATCACTATCAACGAAGGCAAAAAAGCCGTAGCGTGAAGTT 494
|||||
QY 602 ACTAATGAAGGCGCTAAATCTTTCATGTGGGTAGCCATTTCCACCTCTTTTGAAGCTAAC 661
|||||
Db 495 AAAAATGTTGGCGACAGACCGGTTCAAATCGCGTTCACACTTCCATTTCTTTGAAGTGA 554
|||||
QY 662 AAGGCACATAAAATTCGATCGTAAAAAGCCTATGGCAAAACGCTAGATATTCCCTCTGGC 721
|||||
Db 555 AGATGCTTAGACTTTGACAGAGAAAAAACTTTGCGTAAACGCTTAGACATTCGAGCGGG 614
|||||
QY 722 AACACGCTACGCATTTGGGCGAGACAAACCCGCAAGTGCAGTTGATTCCTCTTTGGTGGC 781
|||||
Db 615 ACAGCGGTGAAGATTTCAGCGCTGGCGAAGAAAAATCCGTAGAAATTGATTGACATTTGGCGGT 674
|||||
QY 782 AGTAAAAAAGTGAATTCGCATGAACGGGCTTGTGAATAACATCGCGGATGAACGCCATAA 841
|||||
Db 675 AACAGAAGAATCTTTGGATTTAAACGCAATTTGTTGATAGACAGCAGACAAACGAAAGCAA 734
|||||
QY 842 CATAAAGCGCTTGACAAGGCGAAATCTCAGCGGATTT----- 877
|||||
Db 735 AAAATGCTTTACACAGAGCTAAAGAGCGTGGTTCATGGCGCTAAAGCGATGACAC 794
|||||
QY 878 -----ATCAAGTAAGGAGACTCCCATGAAA---ATGAAAAAACAGAAATATGTAAT 926
|||||
Db 795 TATGTAAAAACAATTAAGGAGTAAAGAAATGAAAAAGATTAGCAGAAAAAGAAATATGTTCT 854
|||||
QY 927 ACCTACGACCCCAACAAAGCGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGAA 986
|||||
Db 855 ATGTATGGTCTACTACAGCGGATAAAGTGAGATTGGCGGATACAGACTTGTATCGCTGAA 914
|||||
QY 987 GTAGAACATGACTATATACCCTATGCGAAGAACTTAAATTTTGGCGGCGGTAAACTATC 1046
|||||

```
Db 915 GTAGAACATGACATACACCATTTATGGCGAAGAGCTTAAATTCGGTGGCGGTAAACCCCTA 974
Qy 1047 CGTGAGGGTATGGGTGAGACCAATAGCCCTGATGAAACACCCCTAGATTTAGTCATCACT 1106
Db 975 AGAAGAGCATAGCCNAATCTACACCCCTAGCAAGAGTGGATTTAATATCACT 1034
Qy 1107 AACCGATGATATCGACTACACCGGATTTACAAAGCCGACATTTGGATTTAAAGCGC 1166
Db 1035 AACGCTTTAATCGTGGATTTACACCGGTATTTATAAAGCGGATATTTGGTATTTAAAGATGC 1094
Qy 1167 ABAATCATGGCATTCGCAAGGAGGAGAAACAAGGACATGCAAGTGGCGTAAAGCCCTCAT 1226
Db 1095 ABAATCGCTGGCATTTGGTAAGGCGGTAAACAAGACATGCAAGATGGCGTTAAAGAACAAT 1154
Qy 1227 ATGCTCGTGGGTGGGCGACAGAACACTAGCAGGGGAAAGGTATGATTTATACCGCTGGG 1286
Db 1155 CTTAGCGTAGGTCCTGCTACTGAAGCCTTAGCCGGTGAAGGTTGATCGTAAACGGCTGGT 1214
Qy 1287 GGAATCGATTTACACACCCACTTCTTTCTCCACAAATTCCTACCGCTCTAGCCAAT 1346
Db 1215 GGTATTTGACACACATCCACTTCAATTCACCCCAACAAATCCCTACAGCTTTTGCAGC 1274
Qy 1347 GCGTTACACCACTGTTTGGAGGGGCGACAGGTCCTGTAGATGSCACGAATGCGACTACT 1406
Db 1275 GGTGTAAACACCACTGATTTGGTGGGAAACCGGTCCTGCTGATGGCACATAATGCGACTACT 1334
Qy 1407 ATCACTCCGGGCAAAATGGAATTTGCACCGCATGTTGGCGGAGAGAGTATTTCTATG 1466
Db 1335 ATCACTCCAGCAGAGAAATTTAAATGATGCTCAGAGGGGCTGAAGATATTTCTATG 1394
Qy 1467 AATGTGGGCTTTTGGGCAAGGCAATPAGCTCTAGCAAAAACAACTTTGTAGAACAAAGTA 1526
Db 1395 AATTTAGGTTTCTTGGCTAAAGGTAACGCTTCTTAACGATGCGAGCTTTAGCCGATCAAAAT 1454
Qy 1527 GAAGCGGGCGGATGGTTTTAAATTCGATGAAGACTGGGGCGACAAACCAAGTGGCGATC 1586
Db 1455 GAAGCGGGTGCATTTGGCTTTAAATTCAGGAAGACTGGGGCACCACTCTCTTCTGCAATC 1514
Qy 1587 GATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACCGGATACA 1646
Db 1515 AATCATCGGTTAGATGTTGGGCAAAATACGATGTGCAAGTGCATATCCACAGACACT 1574
Qy 1647 GTCAATGAGGCGGTTATGATGACACCCCTAAATGCAATGAACGGGCGCGCCATCCAT 1706
Db 1575 TTGAATGAAGCGGTTGTTAGAGACACTATGCTGCTATTTGCTGGACGCACTATGCAC 1634
Qy 1707 GCCTACACATTTAGGAGCGGTTGGAGGACACTCACTGATGTTATCACCATGGCAGGC 1766
Db 1635 ACTTTCCACACTGAAGCGGCTGGCGGGGACACGCTCTGATATATTAAAGTAGCCGGT 1694
Qy 1767 GAGCTCAATTTCTACCTCTCCACACCCCACTATTCCTATACCATTAATACGGTT 1826
Db 1695 GAACACAACTCTTCCCGGCTTCCACTAAACCCCACTCCCTTTCCCGTGAATACAGAA 1754
Qy 1827 GCAGAACACTTTAGACATGCTCATGACATGCCACCACTTAGACAAACGCAATCCGCGAGGAT 1886
Db 1755 GCAGAGCACATGGACATGCTTATGTTGTGCCACCACTTGGATAAAGCATTAAGAAGAT 1814
Qy 1887 TTACAAATTTCTAAAGCGTATCCGCGCCGGCTCTATTCGGGCTGAAGATGTGCTCCAT 1946
Db 1815 GTTCAGTTCGCTGATTTCAAGGATCCGCGCTCAAAACCAATTTGCGGCTGAAGACACTTTGCAT 1874
Qy 1947 GATATGGTGTGATCGGATGACAGCTCGGATTCGCAAGCAATGGGGCGTGCAGGCGAA 2006
Db 1875 GACATGGGATTTTCTCAATCACCAGTTCTCACTCTCAAGGCGATGGGCGGTGGGTGAA 1934
Qy 2007 GTGATTCCTCGAACTTGGCAGACTGCGGATGAAGATAAAGAAATTTGGTAAAGCTTCCT 2066
Db 1935 GTTATCACTAGAACTTTGGCAACAGCTGACAAAACAGAAAGAAATTTGGCGGCTTGAA 1994
Qy 2067 GAAGATGGCAAGATAACGATAAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTATC 2126
|||||
```

```
Db 1995 GAAGAAAAAGCGGATTAACGACAACTTCAGGATCAACCGCTACTTGTCTAAATACACCAATT 2054
Qy 2127 AACCCCGCTTTTGACCCACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGATC 2186
Db 2055 AACCCAGCGATCGCTCATGGATTTAGCGAGTATGTAGGTTTCAGTAGAAGTGGGCAAGTG 2114
Qy 2187 GCGGACTTGGTGGTGGGAATCTGCGCTTTTGGCGGTAAACCCAAATCGTGATCAAA 2246
Db 2115 GCTGACTTGGTATTTGTGGAGTCCAGCAATCTTTGGCGGTGAACCCCAACATGATCATCAAA 2174
Qy 2247 GCGGATATGTGGTTCCTCTGAAATGGGCGATTCTAACGGCTGTGTGCCCACTCCCCAA 2306
Db 2175 GCGGATTCATTTGCGTTAAGCCHAAATGGGCGATCGGAACGTTCTATCCCTACCCCAAA 2234
Qy 2307 CCGGTTTATTAACCGCAAAATGTTTGGGCATCACGGCAAGGCGAAATTTGACACAGCATC 2366
Db 2235 CCGGTTTATTAACGAGAAATGTTTCGCTCATCATGTTAAAGCTAAATACGATGCAAAACATC 2294
Qy 2367 ACTTTGTTTCCAAAGTCGCTATGAAATGGGCGTGAAGAAAGCTGGGCTTAGAGCGC 2426
Db 2295 ACTTTGTTGTTCTAAGCGGCTTATGACAAAGGCAATTAAGAAAGAAATTAGGACTTTGAAAGA 2354
Qy 2427 CAAGTTCTACCGGTCAAAAACTGCCGTAAACATCACCAAGAAAGACTTCAAGTTCAACGAC 2486
Db 2355 CAAGTTTGGCGTAAAAAAATTTGCAGAATATCACTAAAGAAAGCATGCCAATTCACGAC 2414
Qy 2487 AAAACGGCAAAAAATCACCGTCGATCCGAAACCTTCGAGGTTTGTAGATGGCAAACTC 2546
Db 2415 ACTACTGCTCACATTTGAAGTCAATCCTGAAACTTTACCATGTGTTTCGTGGATGGCAAGAA 2474
Qy 2547 TGCACCTCTAAACCCACTCGCAAGTGCCTCTAGCCAGCGCTACACTTTCTCTAGG 2604
Db 2475 GTAACCTCTAAACAGCAATAAAGTGAGCTTGGCGCAACTCTTTAGCAATTTCTTAGG 2532

RESULT 8
AAQ90180
ID AAQ90180 standard; DNA; 2619 BP.
XX
AC AAQ90180;
XX
AC
XX
DT 25-MAR-2003 (revised)
DT 02-NOV-1995 (first entry)
XX
XX Helicobacter felis urease ureA/ureB operon.
KW Urease; UreA; UreB; vaccine; Helicobacter infection; UreI;
KW heat shock protein; ss.
XX
XX Helicobacter felis.
XX Location/Qualifiers
FH Key 31..37
FH RBS /*tag= b
FH /*note= "ureA Shine-Dalgarno site"
FT CDS 43..756
FT /*tag= a
FT /*EC number= "3.5.1.5"
FT /*note= "UreA"
FT RBS 756..759
FT /*tag= d
FT /*note= "ureB Shine-Dalgarno site"
FT CDS 766..2475
FT /*tag= c
FT /*EC number= "3.5.1.5"
FT /*note= "UreB"
XX
XX WO9514093-A1.
XX
XX 26-MAY-1995.
XX
XX 19-NOV-1993; 93WO-EP003259.
XX
```

PR	19-NOV-1993;	93WO-EP003259.	
XX	(INSP) INST PASTEUR.		
PA	(INRM) INST NAT SANTE & RECH MEDICALE.		
XX			
XX	Labigne A, Suerbaum S, Ferrero R;		
PI			
XX	WPI; 1995-200383/26.		
DR	P-PSDB; AAR74336, AAR74337.		
XX			
XX	Immunogenic composition against Helicobacter infection - also gene		
PT	fragment(s) and protein(s) from Helicobacter urease gene cluster and heat		
PT	shock protein(s).		
XX			
PS	Claim 10; Fig 3; 128pp; English.		
XX			
CC	The sequence encodes urease UreA and UreB proteins, which are components		
CC	of a novel immunogenic composition capable of inducing protective		
CC	antibodies against Helicobacter infection. The composition may include		
CC	the UreA or UreB proteins, a urease- associated heat shock protein		
CC	(AAR74338-39) or the UreI protein (AAR74340). The composition is used to		
CC	prepare a vaccine for humans or animals, especially against H. pylori and		
CC	H. felis. Antibodies against the proteins may be used for treating		
CC	Helicobacter infection, and primers/probes to the DNA sequence may be		
CC	used for detection of Helicobacter infection. (Updated on 25-MAR-2003 to		
CC	correct PN field.)		
XX			
SQ	Sequence 2619 BP; 772 A; 569 C; 665 G; 613 T; 0 U; 0 Other;		
Query Match 36.9%; Score 1062.6; DB 2; Length 2619;			
Best Local Similarity 65.5%; Pred. No. 2.3e-228;			
Matches 1606; Conservative 0; Mismatches 804; Indels 41; Gaps 2;			
QY	192	ACAAGAGTAATAGGTGAACCTCACACCCAAAGAGCGAGAGGCTTAAAGCTCAACCAACCGCGAAGCCA	311
DB	29	ATAAGAGTTTAGGATGAACCTAACCGCTAAAGAACTAGACAAAGTTAATGCTCCATATG	88
QY	252	CGGGCGAAGTGGTAGAAGCGCAAGCAGAGAGGCTTAAAGCTCAACCAACCGCGAAGCCA	311
DB	89	CGGCAGATTGGCAGAGAAGCGTTGGCGGTGTGAACTCAATATACCGAAGCGG	148
QY	312	TTGCTTACATTAGTGCCTATATATGAGCAGAGCGCGCGTGGAAAAAOCGTTGCC	371
DB	149	TCGCGCTATTAGCGGCGTGTGATGAAAAGCGCGTGTGATGTAATAAAGCGTGGCG	208
QY	372	AGCTTATGAAGTGCATGCCTTTTTCAAAAAGATCAAGTAATGCCCGGGTGGTA	431
DB	209	ATTGTATGCAAGAGCGAGCTTGGCTTTAAAAAGAAAAATGTGATGGACGGGTAGCAA	268
QY	432	ATATGGTCCCGATCTAGGTGTAGAGCCACCTTTCTCGATGGTACGAACTTGTAACTG	491
DB	269	GCATGATTCATGAAGTGGGATGTGAAGCTTAATCCCGATGGAACCAAGCTTGTNACTA	328
QY	492	TGAATTGGCCCATCGAACAGATGAGCACTTCAAAGCGGCGAAGTGAATTTGGTTGCG	551
DB	329	TCACACTCCGGTAGAGGATAATGGCAATTAGCCCCCGCGAGTCTTCTTAAA--AA	385
QY	552	ATAAGACATCGAGCTCAATGACGAGCAAGAGATGACCGAATCGAGTTACTAATGAAG	611
DB	386	ATGAGGACATTACTATTAAACGCGGCAAGAGAGCCATTAGCTTGAAGTGAATAAAG	445
QY	612	GGCTTAATCTTGCATGTGGGTAGCCATTTCCACTCTTTGAAGCTACAAAGGCACTAA	671
DB	446	GGATGCTCTGTGCAAGTGGGATCAATTTCCACTTTCTCGAAGTGAATAAGCTCTTGG	505
QY	672	AAATTCATCGTGAAGAAAGCTTATGGCAACGGCTAGATATCCCTCTTGGCAACACGCTAC	731
DB	506	ACTTCATCGCAAAAAGCTTTTGAACAGCGCTAGACATTGCATCTGAAACAGCGTGC	565
QY	732	GCATTGGGCGAGCAAAACCGCAAGTGCATGTGATTCCTCTTGGTGGCAGTAAAAAG	791
DB	566	GCTTTGAACCCGGGAGGAAAAAGTGTGGAAGCTCATTTGACATCGCGGGAATAAGCGCA	625
QY	792	TGATTGGCATGAACGGGCTTGTGAATTAACATCGCGGATGAACGCCATATAAATGAAGCGC	851
DB	626	TCTATGGCTTTAATTTCTTTGGTGGATCGCAAGCGGATGCCGATGGTAAAAAACTCGGCT	685
QY	852	TTGACAAGGCGAAATCTCACGGATTT-----877	
DB	686	TAAACCGCTTAAAGAAAAAGGTTTTTGGGTCTGTAAACTCGGGTTGTGAAGCGACTAAAG	745
QY	878	----ATCAAGTAAGGAGACTCCCATGAAAAATGAAAAACAAGAAATATGTAAATACCTACG	933
DB	746	ATAAACCAATAAGGAAAAACCATGAAAAAGATTTACGAAAAAGAAATATGTTTCTATGATG	805
QY	934	GACCCACAAAAGCGGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAAGTAGAAC	993
DB	806	GTCCCACTACCGGGATCGTGTAGACTCGCGGACACTGATTTGATCTTGAAGTGGAGC	865
QY	994	ATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGGGTAAAACTATCGTGAAG	1053
DB	866	ATGATTGCACCACTTATGGTGAAGAGATCAAAATTTGGGGCGGTAAAACTATCGTGAATG	925
QY	1054	GTATGGGTGAGAGCAATAGCCCTTGATGAAAAACCCCTAGATTTTAGTCATCACTAACCGGA	1113
DB	926	GGATGAGTCAACCAATAGCCCTAGCTCTTATGAATTTAGATTTGGTCTCACTAACGCC	985
QY	1114	TGATTTATCGACTACACCGGATTTTCAAGCCGACATTTGGAATTAATAAAGCGCAAAATCC	1173
DB	986	TCATTTGGAATATACGGGCAATTTTCAAAAGCCGACATTTGGAATTAATAAAGCGCAAGATTG	1045
QY	1174	ATGGCATTTGGCAGGCGAAGAAACAAGACATCAAGATGGCGTAAAGCCCTCATATGTCG	1233
DB	1046	CAGGCATTTGGCAAGGCGGCAATGAAGACATGAAGATGGCGTAGATAAATCTTTGCG	1105
QY	1234	TGGGTGTGGGCACAGAGCACTAGCAGGGGAAGTATGATTTATCCCGCTGGGGGAATCG	1293
DB	1106	TAGTCTCTGTACAGAGCTTTTGGAGCTGAGGCTTGATTTGAAACCGCTGGTGGCATCG	1165
QY	1294	ATTACACACCCACTTCTTTCTCCACAACAAATTCCTACCGCTTAGCCAAATGGCGTTA	1353
DB	1166	ATACGCATATTCACTTTATCTCTCCCAACAAATCCCTACTGCTTTTGGCAGCGGGTTA	1225
QY	1354	CAACCATTTTGGAGGCGGACAGGTCTCTAGATGGCAGCAATGGCACTACTATCACTC	1413
DB	1226	CAACCATGATTTGGAGGCGGACAGGACCTCGGATGGCAGCAATGGCACCACCACTACTC	1285
QY	1414	CGGGCAATTTGCAACCGCATTTTGGCGCAGCAGAGAGATTTCTTATGAATGTGG	1473
DB	1286	CGGACCGCGTAACTCTAAAGATGTTTGGCTGAGCGGCAAGAAATAGCCCATGAATCTAG	1345
QY	1474	GCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTTGTAGAAACAAGTAGAAGCGG	1533
DB	1346	GCTTTTGGCTAAGGGGAATGTGCTTTACGAAACCTCTTTACGCGATCAGATTGAAGCAG	1405
QY	1534	GGCGAATTTGTTTAAATTTGCATGAAGACTGGGGGCAACACCAAGTGGCATCGATCACT	1593
DB	1406	GGCGCATTTGTTTAAATTTCCAGAGACTGGGGGAAGCACACCTGCGAGCTATTCCACACT	1465
QY	1594	GCTTGAGCTGCGAGATGAATACGATGTCAAGTTTGTATCCACACCGATACAGTCAATG	1653
DB	1466	GCCTCAATGTGCGCGATGAATACGATGTGCAAGTGGCTATCCACACCGATACCTTAAAG	1525
QY	1654	AGCGAGTTTATGTAGATGACACCCCTAAATGAACGCGGCGGCCCATCCATGCTTACC	1713
DB	1526	AGCGGGCTGTGTAGAAGACACCTTAGAGCGGATTTGGGGGCGCACCATTCATACCTTCC	1585
QY	1714	ACATTGAGGAGCGGTGAGGACACTCACTCAATGTTTATCACCATGGCAGGCGAGCTCA	1773
DB	1586	ACACTGAAGGGGTGGGGGTGACAGCTCCAGATGTTATCAAAAATGGCAGGGGAATTTA	1645
QY	1774	ATATTCTACCTCTCCACACCCCACTATTTCCCTATATCACTTAATACGTTGTCAGAAC	1833
DB	1646	ACATTTCTACCCGCTCTACTTAACCCGACCAATCTCTTTTCCAAAAACACCTGAAGCCGAGC	1705
QY	1834	ACTTAGCATGCTCATGACATGCCACCACTAGACAAACGCAATCCGCGAGGATTTACAAT	1893


```
Db 1706 ACATGACATGTTAATGGTGGCCACCTGTGATAAAAGTATCAAGGAAGATGTGAGT 1765
Qy 1894 TTTCTCAAAAGCCGTATCCGCGCGCTCTATCGCGCTGAAGATGTCTCCATGATGG 1953
Db 1766 TTGCCGATTCGAGGATTCGCCCCCAACTACTCGCGCTGAGACCAACTCCATGATGG 1825
Qy 1954 GTGTGATCGCGATGACAAGCTCGGATTCGCAAGCAATGGGCGTGCAGGGCAAGTATTC 2013
Db 1826 GGATCTTTTCTATACACAGCTCGGACTCTCAGGCTATGGGACGCTAGGCGAGGTATCA 1885
Qy 2014 CTCGAACTTGGCAGATCGCGATGAAGATTAAGAAAGAAATTTGGTAACTTCTCGAGATG 2073
Db 1886 CACGCACTTGGCAGACAGCAGACAAACAAAGAGATTTGGCGCTTGAAGAGAGAA 1945
Qy 2074 GCAAGATTAACGATAATTTCCGATTAAGCGCTACATCTCAAAATACACTATCAACCCCG 2133
Db 1946 AAGCGGATAACGACACTTCCGATCAACAGCTACATCTCTAATATACACCATCAACCCCG 2005
Qy 2134 CTTTGACCCACGCGGTAGGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCGACT 2193
Db 2006 GGATCGCGATCGGATTTCTGACTATGTGGCTCTGTGGAAGTGGGCAAAATACGCCGACC 2065
Qy 2194 TGGTGGTGTGGAATCTGCTTTTGGGCTAAACCCAAATCGTGATCAAGGGCGTA 2253
Db 2066 TCGTGTCTTGGAGTCGGGCTTTCTTTGGCATTAAAGCCCAATATGATTTAAGGGCGGAT 2125
Qy 2254 TGGTGGTCTTCTGAAATGGGCGATTTCTAACCGCTCTGTGCCACTTCCCAACCGGTTT 2313
Db 2126 TTAATGGCTCTCTCAATGGGCGATGCCAATGGCTATTTCCACCCCTCAGCGCGTCT 2185
Qy 2314 ATTACCGCGAAATGTGTGGCATCACGGCAAGCGGAAATTTGACACGAGATCACTTTTG 2373
Db 2186 ATTACCGTGAATGTGTGGACACCATCGGAAACAAATTCGACACCAATATCACTTTTCG 2245
Qy 2374 TTTCCAAAGTCGCTATGAATAATGGCTGAAAGAAAGCTGGGCTTAGACGGCGCAAGTTC 2433
Db 2246 TGTCCCAAGCGGCTTCAAGCGAGGGATCAAGAAGAACTAGGGCTAGATCGCGCGCAC 2305
Qy 2434 TACCGGTCAAAATCTGCGTAACATCACCAAGAAAGACTTTCAAGTTTCAACGACAAACGG 2493
Db 2306 CGCAGTGAATAACTGTGCAATATCACTAAGAGACCTCAATTCACGATGTGACCG 2365
Qy 2494 CAAAAATCACGTCGATCCGAAACCTTTGAGGTCTTTGTAGATGGCAAACTCTGCACTT 2553
Db 2366 CACATATTGATGTCAACCTGAAACCTATAAGTGAAGTGAAGTGGATGGCAAGAGTAACTT 2425
Qy 2554 CTAAACCCACTCGCAAGTGCCTCTAGCCAGCGCTACACTTTCTTCTAGG 2604
Db 2426 CTAAAGCAGCAGATGAATTGAGGCTTAGCGCAACTTTTATAATTTGTTCTAGG 2476
```

```
RESULT 9
AAT45680
ID AAT45680 standard; DNA; 2619 BP.
XX
AC AAT45680;
XX
DT 16-OCT-2003 (revised)
DT 07-MAR-1997 (first entry)
XX
XX H. felis structural urease gene.
XX
KW Urease; UreA; UreB; heat shock protein A; HspA; vaccine; immunogen;
KW antigen; antibody; chronic gastritis; ulcer; pIL1205; ss.
XX
OS Helicobacter felis; strain ATCC 49179.
XX
XX Key Location/Qualifiers
XX RBS 31..36
XX CDS 43..753
XX /*tag= a
XX /*tag= b

/standard_name= "URE A"
/product= "urease subunit A"
756..759
/*tag= c
766..2475
/*tag= d
/standard_name= "URE B"
/product= "urease B subunit"
WO9634624-A1.
XX
XX 07-NOV-1996.
XX
XX 02-MAY-1996; 96WO-EP001834.
XX
XX 19-MAY-1995; 95US-00432697.
XX
XX 19-MAY-1995; 95US-00447177.
XX
XX (INSP ) INST PASTEUR.
XX
XX (INRM ) INST NAT SANTE & RECH MEDICALE.
XX
XX Labigne A, Suerbaum S, Ferrero RL, Thiberge J;
XX
XX WPI; 1996-505900/50.
XX
XX P-PSDB; AAW06729, AAW06730.
XX
XX New immunogenic compsn. contg. UreB and HspA antigens of Helicobacter -
XX for treatment and prevention of esp. H pylori infection, also new
XX antibodies specific for these antigens.
XX
XX Example 1; Page 96-100; 184pp; English.
XX
XX A portion (AAT45680) of plasmid pIL1205 (NCIM I-1355) comprises the
XX Helicobacter felis structural urease gene encoding urease subunits A
XX (AAW06729) and B (AAW06730). pIL1205 was obtd. by examining the
XX expression of urease activity by H. felis cosmid clones and
XX identification of the H. felis genes required for urease expression when
XX cloned in E. coli. The urease gene can be used to design probes for the
XX detection of Helicobacter infection, or utilised in the prodn. of
XX recombinant urease subunits for use in novel immunogenic compsns.
XX (Updated on 16-OCT-2003 to standardise OS field)
XX
XX Sequence 2619 BP; 772 A; 569 C; 665 G; 613 T; 0 U; 0 Other;
XX
Query Match 36.9%; Score 1062.6; DB 2; Length 2619;
Best Local Similarity 65.5%; Pred. No. 2.3e-238;
Matches 1606; Conservative 0; Mismatches 804; Indels 41; Gaps 2;

Qy 192 ACAAGAGTAAATAGTGAAGTCAACCCAAAGAGCAAGTTCCTTTATATATG 251
Db 29 ATAAGGAGTTAGGATGAACCTAACGCTTAAAGAACTAGACAAGTATGCTCCATTATG 88
Qy 252 CGGCGAAGTGGCTAGAAAGCGCAAGAGGCTTAAAGCTCAACCAACCGGAAGCCA 311
Db 89 CGGCGAATTGGCAGAAAGAACGCTTGGCGCTGTGTGAAGAACTCAATTACACCGAAGCGG 148
Qy 312 TTGCTTACATTAGTGCCTCATATTATGACGAAGCGCGCTGGAAAAAACCCTTTGCC 371
Db 149 TCGCGCTCATTTAGCGGGCTGTGATGAAAGGCGCGTGTATTAAGACGCTGGCGG 208
Qy 372 AGCTTATGGAAGAGTGCATGCATCTTTTGAAGAAAGATGAAGTAAATCCCGGGGGTGA 431
Db 209 ATTTGATGCAAGAGGCGAGGACTTTGGCTTAAAAAAGAAATGTGATGACGCGTAGCAA 268
Qy 432 ATATGTTCCCGATCTAGGTGTAAGCCACTTTCTGTGATGGTACGAACTTGTAACTG 491
Db 269 GCATGATTCATGAAGTGGGATTTGAAGCTAACTTCCCGCATGAAGTGAACCAAGCTTGA 328
Qy 492 TGAATTGGCCCATCGAACCCAGATGACACTTTCAAGCGGGCGAAGTGAATTTGGTTGG 551
Db 329 TCCACACTCCGGTAGAGGATTAATGGCAATTAGCCCCCGCGGAGGTCTTCTTAA--NA 385
Qy 552 ATAAAGACATCGAGCTCAATGCAGGCAAGAAAGTAACCGAACTTGAGGTTACTAATGAAG 611
```

Db 386 ATGAGACATTACTATTAAACGCGGCAAGAGCCATTAGCTTTGAAGTGAATAAAG 445
Qy 612 GGCCTAAATCCTTGCAATGGGTAGGCATTTCCACATCTTTGAAGCTAAACAGGCACTAA 671
Db 446 GGCATCGTCTGTGCAGGTGGGATCAATTTCCATCTTTTCGAAGTGAATAGCTTTGG 505
Qy 672 AATTCGATCGTGAAGAGCTTAGGCAACCGCTAGATATCCCTCTGGCAACACGCTAC 731
Db 506 ACTTCGATCGCAAAAAGCTTTTGCAACGCTAGACATTCGATCTGGAACAGCGGTGC 565
Qy 732 GCATTTGGGCGAGCAAAACCGCAAGTGCAGTTGATCTCTTGTGTGCGAGTAAAGAG 791
Db 566 GCTTTGAACCCGGGAGGAGAAAGTGTGGAACCTCATTTGACATCGCGGGGAATAAGCGCA 625
Qy 792 TGAATGGCATGAACGGCTTGTGAATAACATCGCGGATGAACGCCATAAACAAGCGC 851
Db 626 TCTATGGCTTTAATCTTTTGGTGGATCGCAAGCCGATCCGATGGTAAAAAATCTCGGCT 685
Qy 852 TTGACAAGCGGAAATCTCAGGATTT----- 877
Db 686 TAAACGCGCTAAGAAAAGGTTTGGGTCTGTAAACTGCGGTTGTGAAGCGACTAAG 745
Qy 878 ----ATCAAGTAAAGGACTCCCATGAAATGAAAAAACAAGATATGTAATACCTAG 933
Db 746 ATAAACAAATAAGGAAAAACCATGAAAAAGATTTACGAAAAAGATATGTTTCTATGATG 805
Qy 934 GACCCACCAAGGCGATAAGTGCCTTAGGAGATACCGATCTTTGGGCGAGAGTAGAAC 993
Db 806 GTCCCACTACCGGGGATCGTTGTAGACTCGCGACACTGATTTGATCTTTAGAAATGGAGC 865
Qy 994 ATGACTATACCACTATGCGGAAGAACTTAAATTTGGCGCGGTAAAACTATCCGTGAGG 1053
Db 866 ATGATTGACCACTTATGTGGAAGATCAAAATTTGGGGCGGTAAAACTATCCGTGATG 925
Qy 1054 GTATGGGTGAGAGCAATAGCCCTGATGAAAAACACCTAGATTAGTTCATCAACGCGCA 1113
Db 926 GGATGAGTCAAAACCAATAGCCCTAGCTCTTATGAATTAGATTGGTGTCTCACTAAGCC 985
Qy 1114 TGAATTATGACTACACCGGGATTTCAAGCGGACATTTGGGATTTAAAAACGCAAAATCC 1173
Db 986 TCAATTTGAGCTATACGGGCAATTTCAAGCGGACATTTGGGATTTAAAGACGCAAGATTG 1045
Qy 1174 ATGGCATTTGGCAAGCAGCAACAGGACATCGAGATGCGCTTAAGCCCTCATATGGTGC 1233
Db 1046 CAGGCATTTGGCAAGGAGGCAATAGGACATGCAAGATGGCTGAGTAATAATCTTTTGGC 1105
Qy 1234 TGGGTGTGGGCACAGAACACTAGCAGGGGAAGGTATGATTATTACCGCTGGGGGAATCG 1293
Db 1106 TAGGTCTCTACAGAGCTTTGGCAGCTGAGGGCTTGATTGTAAACCGCTGTGTGCATCG 1165
Qy 1294 ATTCAACACCCACTTCTCTTTCACAACAATTCCTACCGCTCTAGCCAATGGCGTTA 1353
Db 1166 ATACGCATATTCACTTTATCTCTCCCAACAAATCCCTACTGCTTTTGGCAGCGGGTTA 1225
Qy 1354 CAACCATGTTTGGAGCGGCAAGGTCTGTGATGGCAGCAATGGGACTATATCACTC 1413
Db 1226 CAACCATGATTGGAGGAGGCAAGGACCTTGGGATGGCAGCAATGGGACCATCACTC 1285
Qy 1414 CCGGCAATGGAATTTGCAACCGCATGTTGCGCGCAGCAGAAGATTTCTATGAATGTGG 1473
Db 1286 CCGGAGCGGCTAATCTAAAGAATGTTGCGTGTGAGCCGAGAGATACGCCATGAATCTAG 1345
Qy 1474 GCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATTTAGAACAAAGTAGAGCGG 1533
Db 1346 GCTTTTGGCTTAAGGGAATGTCTTACGAACCTCTTTACCGCATCAGATTGAAGCAG 1405
Qy 1534 GCGCATTTGGTTTTAAATGTCATGAAGACTGGGGCAACAACCAAGTCGATCACT 1593
Db 1406 GGGCGATTGGTTTTAAATCCACGAAGACTGGGGAAGCAACCTCGCAGCTATTCACCACT 1465
Qy 1594 GCTTCAGCGTGCAGATGAATACGATGCAAGTTTGTATCCACCGCATACAGTCAATG 1653

Db 1466 GCCTCAATGTGCGCGATGAATACGATGTGCAAGTGGCTATCCACACGATACCTTTAAG 1525
Qy 1654 AGGCAGTTTATGTAGATGACACACCTAAATGAACGCGCGGCCCATTCATCGCTACC 1713
Db 1526 AGCGGGCTGTGTAGAAGACACCTTAGAGGGATTGCGGGCGCACCATCCATACCTTCC 1585
Qy 1714 ACATTAAGGAGCGGTGTGAGGACACTCACTGATTTATCAACCATGGCAGCGAGCTCA 1773
Db 1586 ACATTAAGGGGCTGGGGTGGACACGCTCCAGATGTTATCAAAATGGCAGGGGAATTTA 1645
Qy 1774 ATATTTCTACCTCTCCACACCCCGCTATTCCTCTATACCATTAATACGTTTCGAGAC 1833
Db 1646 ACATTTCTACCGCTCTACTAACCCGACCATCTCTTTCCAAAAACACTGAAGCCGAGC 1705
Qy 1834 ACTTACATGCTCATGACATGCCACCATAGACAAACGATCCGCGAGGATTTACAAT 1893
Db 1706 ACATGACATGTTAATGGTGTGCCACCATTTGATATAAAGTATCAAGGAAGATGTGAGT 1765
Qy 1894 TTTCTCAAGCCGTATCCGCCCGGCTCTATCGCGGTGAAGATGTGCTCATGATATGG 1953
Db 1766 TTGCCGATTCGAGGATTCGCCCCCAAACTATTCGGCTGAAGACCAACTCCATGACATGG 1825
Qy 1954 GTGTGATCCGATGACAGCTCGGATTCGCAAGCAATGGGCGTGCAGGCGAAGTATTC 2013
Db 1826 GGATCTTTCTATCACCGCTCCGACTCTCAGGCTATGGACGCTAGGCGAGGTATCA 1885
Qy 2014 CTCGAATTTGGCAGACTGCGGATAAGAAATAAAAAAGATTTGGTAAAGCTTCTCAAGATG 2073
Db 1886 CAGCCTTTGGCAGACAGCAGACAAAAAAGAGTTTGGCGCTTGAAGAGGAAA 1945
Qy 2074 GCAAGATAACGATAATTTCCGATTTAAGCGCTACATCTCCAATACACTATCAACCCCG 2133
Db 1946 AAGCGATTAACGACAACTTCGCGATCAACGCTACATCTCTAAATACACCATCAACCCCG 2005
Qy 2134 CTTTGACCAACGCGGTGAGGAGTATTCGCGCTCTGTGGAAGAGGCGAAGATCCCGACT 2193
Db 2006 GGATCGCGATGGGATTTCTGACTATCTGCGCTCTGTGGAAGTGGGCAAAATACGCCGACC 2065
Qy 2194 TGGTGTGTGGAATCTCGCTTTTGGCGTAAACCCAAATCGTGATCAAGAGCGGTA 2253
Db 2066 TCGTGTCTTTGGAGTCCGGCTTCTTTTGGCATTAAGCCCAATATGATTTAAGGGCGGAT 2125
Qy 2254 TGGTGTCTTCTCTGAATGGGCGATTTCTAAACGCTCTGTGCCACTCTCCCAACCGGTTT 2313
Db 2126 TTATTTGGCTCTCTCAATGGGCGATGCCAATGGGTCTATTTCCACCCCTCAGCCGCTCT 2185
Qy 2314 ATTACCGCAAAATGTTTGGGCATCAACGCAAGCGAAATTTGACACAGCATCACTTTTG 2373
Db 2186 ATTACCGTGAATGTTTGGACACCATGGGAAAAAACAATTCGACACCAATATCACTTTTCG 2245
Qy 2374 TTTCCAAAGTCGCTATGAATGGCGTGAAGAAAGCTGGGCTTAGAGCGGCAAGTTC 2433
Db 2246 TGTCCCAACGCGCTTCAAGGCGAGGATCAAGGAAGACTAGGGCTAGATTCGCGCGCAC 2305
Qy 2434 TACCGGTCAAAACCTCCGTAAACATCAACCAAGAAAGACTTCAAGTTTCAACGACAAAAACGG 2493
Db 2306 CCGCAGTGAAGAACTCTGCGAATATCACTAAAAGGACCTCAATTCACGATGTGACCG 2365
Qy 2494 CAAAAATCACCGTTCGATCCGAAAACCTCTCGAGTCTTTGTAGATGGCAAACTCTTCGACCT 2553
Db 2366 CACATATTGATGTCAACCCCTGAAACCTATAAGGTGAAGTGGATGGCAAGAGTAACCT 2425
Qy 2554 CTAAACCACTCTGCAAGTGCCTCTAGCCCGAGCGCTACACTTTCTTCTAGG 2604
Db 2426 CTAAAGCAGCAGATGAATTTAGCGCTAGCGCAACTTTATATATTTTGTCTAGG 2476

RESULT 10
AAQ75319
ID AAQ75319 standard; DNA; 2619 BP.
XX
AC
XX AAQ75319;

CC	sequence, therefore all encoded peptides are cross-referenced to both					
CC	sequences.					
XX						
SQ	Sequence	4824 BP; 1389 A; 1090 C; 1146 G; 1199 T; 0 U; 0 Other;				
	Query Match	36.7%; Score 1058.6; DB 5; Length 4824;				
	Best Local Similarity	65.1%; Pred. No. 2.2e-227;				
	Matches 1606; Conservative	0; Mismatches 797; Indels 38; Gaps 3;				
Qy	197	GAGTAATAGGTGAAACTCACACCAAGACGAAGAAAAGTTCTTGTTATATTATGGGGC	256			
Dd	77	GAGATCTCCATGAAACTCACCCAAAGAGTGTAGATAAGTTGATGCTCCAACAGCTGGA	136			
Qy	257	GAAGTGGCTAGAAGCGCAAAGCAGAGGGCTTAAAGCTCAACCACCCGAAGCCATTGCT	316			
Dd	137	GAATTGGCTAAAAAACGCAAGAAGAAAAGGCATTAAAGCTTAACCTATGTAGAAGCAGTAGCT	196			
Qy	317	TACATTAGTCCCATATTATTGGACGAAGCGCGGTGGAAAAAACAACCGTTGCCAGCTT	376			
Dd	197	TTGATTAGTGCCCAATATTATGGAAGACGAGAGCTGGTAAAAAGACTGCGGCTGAATTG	256			
Qy	377	ATGGAAGAGTGCATGCATCTTTTTGAAAAAAGATGAAGTAATGCCCGGGTGGGTAAATATG	436			
Dd	257	ATGCAAGAGGGCGCACTCTTTTTAAAACCAAGATGATGTGATGATGCGTGGCAAGCATG	316			
Qy	437	GTTCGCCATCTAGGTGTAGAAGCACACTTTCCTGATGGTAGAGAAACTTGTAACTGTGAAT	496			
Dd	317	ATCCATGAAGTGGGTATTTGAAGCGATGTTTCTGATGGGACTAAACTCGTAACCGTGCAT	376			
Qy	497	TGGGCCATCGAACCCAGATGAGCACTTCAAAGCGGCGAAGTGAAATTTGGTTGGCATAAA	556			
Dd	377	ACCCTATTGAGGCCAATGTTAAATTAGTTCCTGGTGATTG---TTCTTAAAAAATGAA	433			
Qy	557	GACATCGAGCTCAATGCAAGCAAAGAAGTAACCGAACTTCGAGTTACTAATGAAGGGCCT	616			
Dd	434	GACATCACTATCAACGAAGGCMAAAAGCCGTTAGCGTGAAAGTTAAAAATGTTGGCGAC	493			
Qy	617	AATCTCTGATGTGGGTAGGCATTTCCACTTCTTTTGAAGCTAACAGGCACATAAANTTC	676			
Dd	494	AGACCGGTTCAAAATPCGGCTCACACTTCCATTTCTTTTGAAGTGAATAGATGCCCTAGACTTT	553			
Qy	677	GATCGTGA AAAAGCCTATGGCAACGCTAGATAATCCCTCTGGCAACACGCTACGCATT	736			
Dd	554	GACAGAAAAAACTTTTCGTTAAACGCTTAGACATTCGAGCGGACAGCGGTAAAGATT	613			
Qy	737	GGGCGACGACAAACCCCGAAAGTGCATGATTCTCTTGGTGGCAGTAAAAAGTGATT	796			
Dd	614	GAGCTGGCGAAGAAAAATCCGTAGAATTGATTGACATTTGGCGGTAAACAGAAGAACTTT	673			
Qy	797	GGCATGAACGGGCTTGTGAATAACATCGCGATGAACGCCATAAACATAAAGCGCTTGAC	856			
Dd	674	GGATTTAACGCATTGGTTGTATAGACAAGCAGACAAACGAAACAAAAAATTTGCTTTACAC	733			
Qy	857	AAGCGCAAAATCTCACGGATT-----ATCAAGT	884			
Dd	734	AGACTAAAGAGCGTGGTTTCATGGCGCTAAAGCGATGACAATATGTAAAAACAATT	793			
Qy	885	AAGAGACTCCCATGAAA---ATGAAAAACAGAATATGTAAATACTACGGAACCCACC	941			
Dd	794	AAGGAGTAAGAAATACAAAAAGATTTAGCAGAAAAAGAAATATGTTTCTATGTATGGTCTACT	853			
Qy	942	AAAGGCGATAAGTCGCTTAGGAGATACCGATCTTTGGCGAGAGTAGBACATGACTAT	1001			
Dd	854	ACAGGCGATAAGTGTAGATTGGGCGATACAGACTTGTATCGTGAAGTAGAACATGACTAC	913			
Qy	1002	ACCACTTATGGCGAAGAACTTTAAATTTGGCGCGGTAAAACTATCCGTGAGGGTATGGGT	1061			
Dd	914	ACCATTTATGGCGAAGACTTAAATTCGTTGGCGGTAAAAACCCCTAAGAGAAGGCATGAGC	973			
Qy	1062	CAGAGCAATAGCCCTGATGAAAAACCCCTAGATTAGTCATCACTAACCGGATGATTATC	1121			
Dd	974	CAATCTAACAAACCTTAGCAAGAAGAGTTGATTTAATTAATCACTAACCGCTTTAATCGTG	1033			

Qy	1122	GA	CTACACACGGGATTTTACAAAGCCGACATTTGGGATTTAAAAACGGCAAAATTCATTGGCATT	1181
Db	1034	GATTACACCGGTATTTATTAAGCCGATATTTGGTATTTAAAGATGGCAAAATCGCTGGCATT	1093	
Qy	1182	GGCAAGCGCAGGAACAAAGGACATGCAAGATGGCGTAAAGCCCTCATATGGTCGTGGGTGTG	1241	
Db	1094	GGTAAAGCGCGTAAACAAAGACATGCAAGATGGCGTTAAAAACAATCTTAGCGTAGGTCTCT	1153	
Qy	1242	GGCACAGAAGCACTAGCAGGGGAAGGTATGATTATTACCGCTGGGGAAATCGATTACAC	1301	
Db	1154	GCTACTGAAGCCTTAGCCGCTGAAGGTTCATCGTAACGGCTGGTGTATTGACACACAC	1213	
Qy	1302	ACCGACTTCCTTCTCCACAAACAATTTCCCTACCGCTCTAGCCAAATGGCGTTTACAAACCATG	1361	
Db	1214	ATCCACTTCAATTTACGCCCAACAAATCCCTACAGCTTTTTCGAAGCGGTGTAAACAACCATG	1273	
Qy	1362	TTTGGAGCGGCACAGGTCTCTGTAGATGGCACGAATGCGACTACTATCACTCCGGGCAAA	1421	
Db	1274	ATTGGTGGTGGAAACCGGTCTCTGTATGGCACATAATGCGACTACTATCACTCCAGGCAGA	1333	
Qy	1422	TGGAACTTGCACCGCATGTTTCGCGCGCAGCAGAAAGATTTCTATGAATGTGGGCTTTTGTG	1481	
Db	1334	AGAAATTTAAATGGATGCTCAGAGCGCTGAAGAAATATTCTATGAATTTAGGTTTCTTGTG	1393	
Qy	1482	GGCAAAAGGCAATAGCTCTACGCAAAAAACAATCTGTGTAGAAACAAGTAGAAGCGGGCGCAT	1541	
Db	1394	GCTAAAGGTAAACGCTTTAAACGATGCGAGCTTAGCCGATCAAAATTGAAGCCGGTGCGCAT	1453	
Qy	1542	GGTTTTAAATTCGATGAAGACTGGGCACACCAACGAATGCGATCGATCACTGCTTGAGC	1601	
Db	1454	GGCTTTGCAATTCACGAAGACTGGGGCACCACTCTCTCTCGAATCAATCATGGGTTAGAT	1513	
Qy	1602	GTGCGAGATCAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATAGGCGCAGGT	1661	
Db	1514	GTTCGGACAAATACGATGTGCAAGTCGCTATCGCCACAGACACTTTTGAATGAAGCCGGT	1573	
Qy	1662	TATGTAGATGACACCCCTAAATGCAATGAAACGGGGCGGCCATCCATGCGCTACCACTTGAG	1721	
Db	1574	TGTGTAGAGACACTATGTGGCTGTATTGTCTGGACGCACTATGACACATTTTCCACACTGAA	1633	
Qy	1722	GGAGCGGTGGAGACACTCACTGATGTTATCACCATGCGACGCGAGCTCAATATTCTA	1781	
Db	1634	GGCGCTGGCGCGGACACGCTCTCTGATATTATTAAGTAGCCGGTGAACACAACTTCTT	1693	
Qy	1782	CCCTCTCTCACACCCCCCACTATTCCCTCTATACCAATTAATACGGTTGCAGAACACTTAGAC	1841	
Db	1694	CCGCTTCCACTAACCCCAACCTCCCTTTACCGTGAATACAGNAGCAGAGCACATGAC	1753	
Qy	1842	ATGCTCATGACATGCCACCACTAGACAAACGCATCCGCGAGGATTTTCAATTTTCTCAA	1901	
Db	1754	ATGCTTATGGTGTGCCACCACTTGGATAAAAGCAATTAAGAAAGATGTTCAGTTCCGTGAT	1813	
Qy	1902	AGCGTATCCGCCCGGCTCTATCCGGCTGAAGATGTGCTCCATGATATGGGTGTGATC	1961	
Db	1814	TCAAGGATCCGCCCTCAAAACCATTTCCGGCTGAAGACACTTTTGCATGACATGGGGATTTTC	1873	
Qy	1962	CGCATGAAGCTCGGATTCGCAAGCAATGGGGGTGCAGCGCAAGTGATTTCCTCGAACT	2021	
Db	1874	TCAATCACAGTTCTGACTCTCAAGCGATGGGCCGTGTGGTGAAGTTATCACTAGAACT	1933	
Qy	2022	TGGCAGACTCGCGATAAGATAAAAAAGAAATTTGGTAAGCTTCTCTGAAGATGGCAAGAT	2081	
Db	1934	TGGCAACAGCTGACAAAAACAAGAAAGAAATTTGGCCGCTTGAAGAAAGAAAAAGGCGAT	1993	
Qy	2082	AACGATAATTTCCGCATTTAAGCGCTACATCTTCAAATATACACTATCAACCCCGCTTTGACC	2141	
Db	1994	AACGACAACTTCAGGATCAAAACGCTACTTGTCTAAATACACCAATTAACCCAGCGCATCGCT	2053	
Qy	2142	CACGGCTGAGCGAGTATATCGGCTCTGTGTGAAGAGGCGCAAGATCGCCGACTTGGTGGTG	2201	
Db	2054	CATGGGATTTAGCGAGTATGTAGGTTCTAGTAGAGTGGGCAAAGTGGCTGACTTGGTATTGTG	2113	
Qy	2202	TGGAATCTGCTCTTTTTTTGGCGTAAAAACCCAAAAATCGTGAATCAAGAGCGGTATGCTGGTG	2261	

Db 2114 TGGAGTCCAGCATCTTTGGCGTGAACCCCAACATGATCATCAAGCGCGATTTCATTGCG 2173
QY 2262 TTCTCTGAAATGGCGATTCTTAACGGCTGTGCGCCACTCCCAACCGGTTTATTACCGC 2321
Db 2174 TTAAGCAATGGCGATGGCAAGCTTCTATCCCTACCCACACCGGTTTATTACAGA 2233
QY 2322 GAAATGTTGGGATCACCGCAAGCGAAATTTGACACGAGCATCTTTTGTTCCTCAA 2381
Db 2234 GAAATGTTGGCTCATCATGTAAGCTAAATACGATGCAACATCACTTTTGTCTCAA 2293
QY 2382 GTGCGCTATGAATGGCTGGAAGAAAGCTGGCTTAGCGCCCAAGTTCTACCGTC 2441
Db 2294 GCGGCTTATGACAAAGGCAATTAAGAAGAAATTAGGACTTGAAGACAGTGTTCGCGGTA 2353
QY 2442 AAAAACTGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGACAAAAACGCAAAATC 2501
Db 2354 AAAAACTGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGACAAAAACGCAAAATC 2413
QY 2502 ACCGTGATCCGAAACCTTCGAGGCTTTTGTAGATGGCAAACTCTGACCTCTAAACCC 2561
Db 2414 GAAGTCAATCCTGAAACTTACCATGTGTTGCGTGGATGCAAGAAAGTAACTTCTAAACCA 2473
QY 2562 ACTGCGAAGTGCCTCTAGCCAGCGCTACACTTTTCTCTAGG 2604
Db 2474 GCGAATAAAGTGAAGTGTGGCGCAACTCTTTAGCAATTTTCTAGG 2516

RESULT 12

ABE55112
ID ABE55112 standard; DNA; 4824 BP.
XX ABE55112;
AC ABE55112;
DT 22-SEP-2005 (first entry)
DE Salmonella vector pHUR3 expressing helicobacter antigens, SEQ ID 1.
KW Circular; pHUR3; ds; vaccine; helicobacter pylori infection; urease;
KW immune stimulation; antibacterial; Immunostimulant; antiulcer;
KW cytostatic; antiinflammatory; gastrointestinal-gen.; gastritis;
KW peptic ulcer; carcinoma.
XX Helicobacter pylori.
OS Salmonella typhimurium.
OS Escherichia coli.
OS Synthetic.
XX WO200132014-A2.
PN 10-MAY-2001.
PD 10-MAY-2001.
XX 01-NOV-2000; 2000WO-US030191.
PF 01-NOV-1999; 99US-00431705.
PR (ORAV-) ORAVAX INC.
XX Kleanthous H, Londono-Arcila P, Freeman D;
PI WPI; 2001-343379/36.
DR P-PSDB; ABE55113, ABE55114, ABE55115, ABE55116, ABE55117, ABE55118.
DR ABE55119, ABE55120, ABE55121, ABE55122, ABE55123, ABE55124, ABE55125,
DR ABE55126, ABE55127, ABE55128, ABE55129, ABE55131, ABE55132, ABE55133,
DR ABE55134, ABE55135, ABE55136, ABE55137, ABE55138, ABE55139, ABE55140,
DR ABE55141.
PT Inducing an immune response against Helicobacter in mammals, useful for
PT treating Helicobacter induced gastroduodenal diseases.
XX Disclosure; SEQ ID NO 1; 63pp; English.
XX The invention relates to inducing an immune response against Helicobacter

CC in a mammal, comprising mucosally administering to the mammal an
CC attenuated Salmonella vector containing a nucleic acid molecule encoding
CC a Helicobacter antigen, and parentally administering to the mammal a
CC Helicobacter antigen. Also included is an attenuated Salmonella vector
CC comprising a nucleic acid molecule encoding a Helicobacter antigen. The
CC Helicobacter antigen is a urease, a urease subunit, or its immunogenic
CC fragment (encoded by the ureA and ureB genes). The mammal is at risk of
CC having but does not have Helicobacter infection or has a Helicobacter
CC infection. The attenuated Salmonella vector further comprises an htrA or
CC nirB promoter. The vector can be used in inducing an immune response
CC against Helicobacter in a mammal. The vector can be used to treat
CC Helicobacter infection. The vector and the method can be used to treat
CC Helicobacter induced gastroduodenal diseases, including acute, chronic or
CC atrophic gastritis, peptic ulcer diseases (e.g. gastric or duodenal
CC ulcers), and gastric carcinoma. Attenuated Salmonella vector pHUR3 was
CC constructed expressing UreA and UreB and including an E. coli AmpR gene.
CC The present sequence represents the pHUR3 plasmid. NOTE: pHUR3 is
CC presented by both ABE55112 and ABE55130, the peptides expressed by
CC pHUR3 are shown in the sequence listing to be split between ABE55112 and
CC ABE55130, yet Figure 4 shows all the peptides being expressed by the one
CC sequence, therefore all encoded peptides are cross-referenced to both
CC sequences.
XX
SQ Sequence 4824 BP; 1389 A; 1090 C; 1146 G; 1199 T; 0 U; 0 Other;
Query Match 36.7%; Score 1058.6; DB 5; Length 4824;
Best Local Similarity 65.7%; Pred. No. 2.2e-227;
Matches 1606; Conservative 0; Mismatches 799; Indels 38; Gaps 3;
QY 197 GAGTAATAGGTGAACTCACACCCAAAGAGCAAGAAAGTTCTTTATATTATGCGGC 256
Db 77 GAGATCTCCATGAAACTCACCCCAAAAGAGTTAGATAAGTTGATGCTCCCTACGCTGA 136
QY 257 GAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGAGCATGTCT 316
Db 137 GAATTGGCTTAAAAACGCAAGAAAGGATTAAGCTTAACTATGTAGAGCAAGTAGCT 196
QY 317 TACATTAGTGCCCATATTATGACGAAGCGCGCGTGGAAAAAACCCTTGCCAGCTT 376
Db 197 TTGATTAGTGCCCATATTATGAGAAAGCGAGAGCTGTTAAAGACTGCGGCTGAATTG 256
QY 377 ATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGTAATGCCCGGGTGGTAAATG 436
Db 257 ATGCAAGAGGGCGCACTCTTTTAAACCAAGATGATGTGATGCGTGGCAAGCATG 316
QY 437 GTTCCCGATCTAGGTGTAGAGCCACTTCTTCGTAGGTGACGAACTTGTAACTGTGAAT 496
Db 317 ATCCATGAAGTGGGTATTGAAGCGATGTTTCTCTGATGGGACTAAACTCGTAACCGTGCAT 376
QY 497 TGGCCCATCGAACCCAGATGAGCACTTCAAAGCGGCGAAGTGAATTTGGTTGCGATAAA 556
Db 377 ACCCTATTGAGGCAATGGTAAATTAGTTCCTGGTGAGTTG---TTCCTTAAAAAATGAA 433
QY 557 GACATCGAGCTCAATGACGAGCAAGAAAGTAACCGAACTTGAGGTACTAATGAAGGCGCT 616
Db 434 GACATCACTATCAACGAAGGCAAAAGCGCTTAGCGTGAAGAGTTAAATAATGTTGGCGAC 493
QY 617 AAATCCTTGCATGTGGGTAGCCATTTCCACTCTTTTGAAGCTAACAGGCACTAAATTC 676
Db 494 AGACCGGTTCAAATCGGCTCACACTTCCATTTCTTTGAAGTGAATAGATGCCTAGACTTT 553
QY 677 GATCGTGAAGAAAGCCTATGCGAAAGCCCTAGATATTCCCTCTGCGACACGCTACGCATT 736
Db 554 GACAGAGAAAAAATCTTCGGTAAACCTTTAGACATTTGACATTGGCGGTAAACAGATCTTT 613
QY 737 GGGGCGAGGACAAACCCGCAAGTGCAGTTGATTTCTCTTGGTGGCAGTAAAAAAGTGATT 796
Db 614 GAGCTTGGCGAAGAAAAATCCCGTAGAATTTGATTGACATTGGCGGTAAACAGAGATCTTT 673
QY 797 GGCATGAACGGGCTTGTGAATTAACATCGCGGATGAACGGCATTAACATAAAGCGCTTGAC 856
Db 674 GGATTTAACGCATTTGGTTGTATAGACAAGCAGACAAAGAAAAAATTTGCTTTTACAC 733

Qy	857	AAGCGGAATCTCACGGATTTT-----ATCAAGT	884
Db	734	AGAGCTAAAAGAGCGTGGTTTTTTCATGTGGCGCTAAAGAGCATGACAACTATGTATAAAAACAATTT	793
Qy	885	AAGGAGACTCCCATGAAA---ATGAAAAACAAAGATATGTAAATACCTACGACGCCACC	941
Db	794	AAGGATAGAAATAGAAAAGATTAGCAGAAAAGAATATGTTCTATGTATGTGCTCTACT	853
Qy	942	AAAGGGCATAAAGTGGCGCTTTAGGAGATACCGATCTTTTGGCGAGAAGTAGAACATGACTAT	1001
Db	854	ACAGGGATAAAGTGAATTTGGCGATACAGACTTGTATCGTCTGAAGTAGAACATGACTAC	913
Qy	1002	ACCACCTATGGCGAAGAACTTAAATTTGGCGCGGTAAAACTATCCGTAGGGTATGGGT	1061
Db	914	ACCATTTATGGCGAAGAGCTTAAATTTGGTGGCGGTAAAAACCCCTAAGAGAAGCATGAGC	973
Qy	1062	CAGAGCAATAGCCCTGATGAAAAACACCTAGATTTTGTGTATCACTAAACCGATGATTATC	1121
Db	974	CAATCTAACACCTTAGCAAGAGAGTTGGATTTTAATTTACTAAACGCTTTAATCGTG	1033
Qy	1122	GACTACACCGGGATTTTACAAAGCGCAATTTGGGATTTAAAAACGGCAAAATCCATGGCAATT	1181
Db	1034	GATTTACACCGGTATTTATAAGCGGATTTGTTATTAAGATGGCAAAATCGTGGCAATT	1093
Qy	1182	GGCAAGCGAGAAACAAGGACATCAAGATGGCGGTAAAGCCCTCATATGTCGTGGGTGTG	1241
Db	1094	GGTAAAGCGCGTAAACAAAGACATCAAGATGGCGTTAAAAAACAATCTTAGCGTAGGTCT	1153
Qy	1242	GGCACAGACACTAGCAGGGGAGGTATGNTTATTACCGCTGGGGAATCGATTACAC	1301
Db	1154	GCTACTGAAGCCTTAGCCGCTGAAGTGTGATCGTAACCGCTGGTGTATTTGACACACAC	1213
Qy	1302	ACCCACTTCTCTTCTCCACAAACAAATCCCTACCGCTCTAGCCAAATGCGTTTAAACACCATG	1361
Db	1214	ATCCACTTCAATTTACCCCGAACAAATCCCTACAGCTTTTTCGACGCGGTGAAACCAATG	1273
Qy	1362	TTTGGAGCGGCAACAGTCTCTGTATAGTGCAACGAATCGCACTACTACTCTCGGGGAAA	1421
Db	1274	ATTGGTGGTGGAAACCGGTCTCTGTGATGGCACTAAATCGCACTACTACTCCAGGCAGA	1333
Qy	1422	TGGNACTTGCACCGCATGTTTGGCGGACGAGAGAGTATTTCTATGAATGTGGGCTTTTGTG	1481
Db	1334	AGAAATTTAAATTGGATGTCTCAGAGCGGCTGAAGAAATATTTCTATGAATTTAGGTTTCTTG	1393
Qy	1482	GGCAAGGCAATAGCTCTAGCAAAAAACAATCTGTAGAAACAAGTAGAAGCGGCGCGATT	1541
Db	1394	GCTTAAGGTAAACGGTCTCTAAGATGCGAGTGTAGCCGATCAAAATTGAAGCGGTGCGATT	1453
Qy	1542	GGTTTTAAATTTGCAATGAAGACTGGGGCAACAACCAAGTCGATCGATCTGCTTGAGC	1601
Db	1454	GGCTTTGCAATTCACGAAGACTGGGCGACCACTCTCTTCTGCAATCAATCATGCTTAGAT	1513
Qy	1602	GTGGCAGATGAATACGATGTGCAAGTTTGTATTCACACCGGATACAGTCAATAGGCGAGGT	1661
Db	1514	GTTTCGACAAAATACGATGTGCAAGTGGCTATTCGCCACAGACACTTTTGAATGAAGCGCGT	1573
Qy	1662	TATGTAGATGACACCTTAAATGCAATGAACGGGCGGCCATCCCATGCTTACCACATTGAG	1721
Db	1574	TGTGTAGAAGACACTATGGCTGTCTATTGCTGGAGCGCATATGACACTTTCCCACTGAA	1633
Qy	1722	GGAGCGGTGGAGACACTACCTGTATGTTATACCACTAGGCGGAGCTCAATATTTCTA	1781
Db	1634	GGCGCTGGCGGCGACACGCTCTGTATTTAATTAAGTAGCGGTGAACACAACTTCTT	1693
Qy	1782	CCCTCTCCACCAACCCCACTATTTCCCTATACCAATTAATACGGTTGACAGAACTTTAGAC	1841
Db	1694	CCCGCTTCCATAAACCCCAACCATCCCTTTTCCACCGTGAATAACAGAAGCAGACATGGAC	1753
Qy	1842	ATGCTCATGACATGCCCACTAGACAAAGCATCCGCGAGATTTTACAAATTTTCTCAA	1901
Db	1754	ATGCTTATGTGTGCCCACTTGATATAAAGCATTAAGAAAGATGTTCAGTTCGCTGAT	1813
Qy	1902	AGCGGTATCCGCCCGGCTCTATCGCGGCTGAAGATGTGCTTCCATGATATGGGTGTGATC	1961

Db	1814	TC	AAGGATCCGCCTCAAAACCAATTCGGCTGGAAGACACTTTGCATGACATGGGGAATTTTC	1873			
Qy	1962	GG	ATGACAAGCTTCGGATTCGCAAGCAATATGGGCGTGCGAGGCGAAGTGATTCCTCGAACT	2021			
Db	1874	TC	ATCACCAGTCTTGACTCTCAAGCGATGGGCGGTGGGTGAAGTTATCATAGAACT	1933			
Qy	2022	TG	CAGACTTCGGATAAGAATAAAAAGAAATTTGGTAACTTCCTCGAAGATGGCAAGAT	2081			
Db	1934	TG	CAAAACAGCTGACAAAAACAAGAAAGAAATTTGGCCGCTTGAAAGAGAAAAAAGCGAT	1993			
Qy	2082	AA	CGATAATTTCCGATTAAGCGCTACATCTCCAATACACTATCAACCCCGCTTTGACC	2141			
Db	1994	AA	CGAACACTTCAGGATCAACCGCTACTTGTCTAAATACACCATTAACCCAGCGATCGCT	2053			
Qy	2142	CAC	GGCGTTCAGCGAGTATATCGCTCTGTGGAAAGGCGCAAGATCGCCGACTTTGGTGGTG	2201			
Db	2054	CAT	GGATTAAGCGAGTATGTAGTTCAAGTAGAGTGGGCAAGTGCTGACTTGGTATTG	2113			
Qy	2202	TG	GAATCTCGCTTTTTCGGCGTAAACCCAAATACGTATCAAGGCGGTATGTGGTGC	2261			
Db	2114	TG	GAGTCCAGCATTTCTTTGGCGTGAAACCCCAACATGATCATCAAGGCGGATTCATTGCG	2173			
Qy	2262	TT	CTCGAAATGGGCGATTTCTAACCGCTCTGTGCCACTTCCCAACCGGTTTATTACCGC	2321			
Db	2174	TT	AAGCCAAATGGCGATCGAAACGCTTCTATCCCTACCCCAACACCGGTTATTACAGA	2233			
Qy	2322	GA	ATGTTTGGGCATCACGGCAAGGCGAAATTTGACACAGCATCACTTTTGTTCGAAA	2381			
Db	2234	GA	ATGTTGCTCATCATCGTAAAGCTAAATACGATGCAAAATCACTTTTGTGCTCAA	2293			
Qy	2382	GT	CCCTATGAAATGGCGTGAAGAAGAAAGCTGGGCTTAGAGCGCCAAAGTTCTACCGTC	2441			
Db	2294	CG	GCCTTATGACAAAGGCATTTAAAGAAAGAAATTAGGACTTTGAAAGACAAGTTTCCCGTA	2353			
Qy	2442	AAAA	CTGCGTAAATCATCACCAAGAAAGACTTCAAGTTCAACGACAAAACGGCAAAATC	2501			
Db	2354	AAAA	ATTGCAGAAATATCATCAAAAAGACATGCAATTTCAACGACACTACCGCTCACTT	2413			
Qy	2502	AC	CGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCC	2561			
Db	2414	GA	AGTCAATCTGAAACTTACCATGTGTCGTGGATGGCAAGTAATCTCTAAACCA	2473			
Qy	2562	AC	CTCGAAGTGCTCTAGCCAGCGCTACATTTCTTCTAGG	2604			
Db	2474	GC	CAATAAAGTGAGCTTGGCGCAACTTTTAGCATTTTCTAGG	2516			
RESULT 13							
AAQ12485							
XX	ID	AAQ12485 standard; DNA; 2767 BP.					
XX	AC	AAQ12485;					
XX	XX	(revised)					
DT	27-AUG-2003	(first entry)					
XX	23-SEP-1991	(first entry)					
DE	DNA encoding A and B subunits of H. pylori urease.						
XX	Gastritis; peptic ulceration; duodenitis; helicobacter; campylobacter; sh.						
KW	Helicobacter pylori.						
XX	OS	Location/Qualifiers					
XX	XX	64..717					
FT	FT	/*tag= a					
FT	FT	/label= subunit A					
FT	FT	/note= "66 kD"					
FT	FT	721..2400					
FT	FT	/*tag= b					
FT	FT	/label= subunit B					
XX	XX	/note= "31 kD"					

PN WO9109049-A.
XX 27-JUN-1991.
XX 19-DEC-1989; 89GB-00028625.
XX 19-DEC-1989; 89GB-00028625.
PR (THRE-) 31 RES EXPL LTD.
XX PA
XX Tabaqchal IS, Clayton CL, Wren BW, Kleanthous H;
PI WPI; 1991-208084/28.
DR P-PSDB; AAR12515, AAR13550.
XX
XX Oligo:nucleotide(s) specific for *Helicobacter pylori* - used as probes and
PT primers to detect *H. pylori* infection, in diagnosis of gastritis, and
PT duodenal and peptic ulceration.
XX
XX Disclosure; Page 16; 28pp; English.
XX
XX The DNA is a 2.7 kb TaqI fragment encoding the A and B subunits of *H.*
CC *pylori* (previously *C. pylori*) urease, i.e. the 66 and 31 kD antigens.
CC From the sequence, probes and primers can be designed for the
CC amplification (by PCR) of the gene, to produce a prods. common to all *H.*
CC *pylori* strains so far tested and which do not occur in other ureases and
CC can therefore be used as a specific indication of the presence of *H.*
CC *pylori*. *H. pylori* (previously *C. pylori*) is strongly implicated in the
CC pathogenesis of gastritis and duodenal and peptic ulceration in man. The
CC primers/probes can be used for the detection of *H. pylori* DNA in gastric
CC mucosa, saliva or faecal samples to provide an early diagnosis of
CC infection. See also AAQ12434-Q12486. (Updated on 27-AUG-2003 to correct
CC OS field.)
XX
XX Sequence 2767 BP; 867 A; 550 C; 635 G; 715 T; 0 U; 0 Other;
SQ

Query Match 36.5%; Score 1053.6; DB 2; Length 2767;
Best Local Similarity 65.7%; Pred. No. 2.5e-226;
Matches 1615; Conservative 0; Mismatches 804; Indels 39; Gaps 4;
182 AAAATTAAACAAGAGTAATAGTGAACCTCACCCCAAGAGAGAGGCTTAAGCTCAACCAA 301
40 ACACCTTTAAGAATAGGAGAAAGAGTGAACCTCACCCCAAGAGAGTGAAGCTTAAGCTTA 159
242 TTATATATAGCGGGAAGTGAAGAGCGAAGCGAAGAGAGAGGCTTAAGCTCAACCAA 301
100 CTCCTACTAGCTGGAGAAATAGCTAAACCAAGCAAGAGAGAGGCTTAAGCTTAAGCTTA 159
302 CCGAAGCCATTTGCTTACATTAAGTGCCTATATATGAGCAAGCGCGCTGGAAAAAAA 361
160 GTGGAGCGGTACGTTGATTAGTGCCTATATATGAGAGAGAGAGAGGCTGGTAAAGAG 219
362 ACGTTGCCAGCTTATGGAAGAGTGCATGCACTTTTGAAGAGAGAGAGTGAAGTGAAGTGC 421
220 ACTGCGGCTGAATATGATGCAAGAGAGGCGCACTCTTTTAAACCGGATGATGTGGAT 279
422 GGGTGGGTAAATATGGTTCCCATCTAGTGTAGAGCCACCTTTCTCTGATGGTAGCAA 481
280 GCGGTGGCAAGCATGATCCATGAAGTGGGTATTTGAAGCGATGTTCTCTGATGGGACCAA 339
482 CTGTGAACCTGTAATGGGCCATCGAACAGATGAGCACTTCAAGCGGCGCAAGTGAAG 541
340 CTCGTACCGTGCATACCCCTATTGAGGCCAATGGTAAATAGTTAGTTCTGTGAGTTG--- 396
542 TTTGGTTGCGATAAAGACATCGAGCTCAAATGAGGCAAGAGAGTAAACCGAATCTTGAGGTT 601
397 TTTCTTAAAAAATGAAGACATCATATCAACGAAGGCAAGAGAGGCGGTAGCGTGAAGTT 456
602 ACTAATGAAGGCGCTTAATCTTGGATGTGGGTAGCCATTTCCATCTTTTGAAGCTAAC 661
457 CCCCCTGTTGGGCGACAGACCGGTTCAATCGGCTCACACTTCCATTTCTTTGAAGTGAAT 516
662 AAGGCACATAAATTCGATCGTGAAGAGGCTATGGCAACGCGCTAGATATTCCCTCTGGC 721

517 AGATGCTTAGACTTTGACAGAGAAAAAACTTTCCGTAAACCGCTTAGACATTTGCGAGCGG 576
722 AACACGCTACGCAATTTGGGCGAGGACAAACCCGCAAGTGCAGTTGATTTCTCTTTGGTGGC 781
577 ACAGCGGTAAAGTTTGAAGCTTCGGAAGAAAAATCCGTAGAGTTGATTCATTTGGCGGT 636
782 AGTAAAAAGTGTATGCGCATGAACCGGCTTTGTGAATAAATCGCGGATGAACGCCATAAA 841
637 AACAGAAAGAAATCTTTGGATTAAACCGGTTGGTGTATAGGCAAGCCGATTAACGAAGCAA 696
842 CATAAAGCGCTTGACAAGCGGAATCTCACGGATTT----- 877
697 AAAATTGCTTTACACAGAGCTAAAGAGCGTGGTTTTCATGCGCGCTAAAGCGATGACAAC 756
878 -----ATCAAGTAAGGAGACTCCCATGAAA---ATGAAAAAAACAAGATAATGTAAAT 926
757 TATGTAAACAATTAAGGAGTAAGAAATGAAAAGATTAGCAGAAAAGATATGCTTCT 816
927 ACCTACGGACCCACCAAGGCGATAAAGTCGCTTAGGAGATACCGATCTTTTGGGCGAAG 986
817 ATGTATGGCCCTACTACAGGCGATAAAGTGAGATTGGGCGATACAGACTTGATCGCTGAA 876
987 GTAGAACATGACTATACCACTATGCGGAAGAACTTAAATTTGGCGGGGTAAACCTATC 1046
877 GTAGAACATGACTATACCACTATTTATGGTGAAGAGCTTAAATTTGGCGGCGGTAAACCCCTA 936
1047 CGTGAAGGTATGGTTCAGAGCAATAGCCCTGATGAACAAACCCCTAGATTTAGTTCATCACT 1106
937 AGAAGAGCATGAGCAATCTAACACCTTAGCAAGAAAGAACTGGATCTATCATCACT 996
1107 AACCGCATGATTTACGACTACACCGGATTTTACAAAGCCGACATTTGGGATTTAAACACGGC 1166
997 AAGCTTTAATCGTGGATTACACCGGTATTTATAAGCGGATTTGTTATTAAGATGGC 1056
1167 AAAATCCATGGCATTTGGCAAGGAGGAAACAAGACATGCAAGATGCGGTAAGCCCTCAT 1226
1057 AAAATCGCTGGCATTTGTTAAAGCGGTAAACAAAGACACGCAAGATGCGGTAAACAAAT 1116
1227 ATGGTCTGGTGTGGGCGACAGAGCACTAGCAGGGGAGAGTATGATTATACCGCTGGG 1286
1117 CTTAGCGTGGGTCTCTGCTACTGAAAGCTTAGCCGGTGAAGGTTGATTGTAACCTGCTGT 1176
1287 GGAATCGATTACACACCCCACTTTCTTCTCAACAAATTCCTACCGCTCTAGCCAAAT 1346
1177 GGTATTGACACACATCCACTTCTCCCCCAACAAATCCCTACAGCTTTTGCAGC 1236
1347 GCGTTTACAAACCATTTTGGAGCGGCGACAGTCTGTAGATGGCAGAAATGCGACTACT 1406
1237 GGTGTAAACAACCATGATTGGTGGCGAACTGGCCCTGCTGATGGCACTAACGCAACCACT 1296
1407 ATCACTCCGGGCAATGGAACTTGCACCGCATGTTGGCGCGACGAGAGAGTATTTCTATG 1466
1297 ATCACTCCAGGTAGAGAAATTTAAAAATTCATGCTCAGAGCGGCTGAAGATATTTCTATG 1356
1467 AATGTGCGCTTTTGGGCAAGCAATAGCTCTAGCAAAAAAACAACCTTTGTAGAACAAAGTA 1526
1357 AACTTTGGTTTCTTGCTAAAGTAAAGCTTCTAACGATCGAAGCTTAGCCGATCAATTT 1416
1527 GAAGCGGGCGGATTTGGTTTAAATTTGCATGAAGACTGGGGCGACAAACCAAGTGCATC 1586
1417 GAAGCTGTGCGATTTGGCTTAAATCCACGAAGACTGGGGCGACCACTCTCTTCTGCAATC 1476
1587 GATCACTGCTTGAGCTGGCAGATGAATAGATGTGGAAGTTGTATCCACACCGATACA 1646
1477 AATCATGCGTTAGATGTTGCGGCAAAATACGATGTGCAAGTTCGCTTCTCCACAGACACT 1536
1647 GTCAATGAGCGAGTTTATGTAGATGACCCCTAAATGCAATGAACGGCGCGCCATCCAT 1706
1537 TTGAATGAAGCGGTTTTCGCTGGAGAGACATATGGGAGCTTATGGCGAGCGCACTATGAC 1596
1707 GCCTACCATTTGAGGAGCGGTTGGAGGACACTCACCTGATGTTATCACCATGGCAGGC 1766

Db 247 AAAGCCGATATTGCAATTAATAATGGCAATTCATGGCAATTTGGCAAGCAGGCAATAAA 306
Qy 1200 GACATGCAAGATGCGGTAAGCCCTCATATGCTGTGGGTGTGGGCACAGAGAAGCACTAGCA 1259
Db 307 GACATGCAAGATGCGGTTTGCAACAATCTTTGGGTGGGCCCTGCTACTAGAGCTTTGGCC 366
Qy 1260 GGGGAAGGTATGATTTATACGCTGGGGGAATGATTCACACCCACCTTCTTCTTCCCA 1319
Db 367 GCTGAAGGGCTGATTTGATCAGCTGGTGGGATTTGACACCCACATCCACTTTATTTCTCCC 426
Qy 1320 CAACAATTCCTTACCGCTCTAGCCAAATGGGTTTACACCATGATTTGGAGGGGCGACAGGT 1379
Db 427 CAACAATTCCTTACCGCTCTAGCCAAATGGGTTTACACCATGATTTGGTGGGGGAACAGGT 486
Qy 1380 CCTGTAGATGCGCAAGATGCGACTACTATCACTCCGGGCAAAATGGAACTTGGCACCGCATG 1439
Db 487 CCAGCTGATGGACTACGCGACTACCATCACTCCGGGGCGCTGGAACCTTTAAACCATG 546
Qy 1440 TTGGCGCAGCAGAAGATTTCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCT 1499
Db 547 CTCGTGCTCTGAAGAATATGCCATGAACCTTGGGCTATTTGGGTAAAGGGAATGTGCT 606
Qy 1500 AGCAAAAAACAATTTAGAACAGTAGAAGCGGGCGGATTTGGTTTAAATTCATGAA 1559
Db 607 TATGAACCTTCCCTGGTGGATCAACTGGAAGCTGGAGCCATTTGGCTTTAAATCCACGAA 666
Qy 1560 GACTGGGGCAACACCAAGTCGATCGATCACTGCTTGAGCGTGGCAGATGAATACGAT 1619
Db 667 GACTGGGTAGCACACCTCGAGCACTACCATTTGCTTGAATGTGGCTGCAATATACGAT 726
Qy 1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGCGAGTTATGTAGATGACACCTTA 1679
Db 727 GTGCAAGTTGCTATCCACACCGATACCTTGAATGAAAGCGGCTGTGTGGAAGACACTTTG 786
Qy 1680 AATGCAATGAAGCGGCGCCATCCATCCCTACACATTTAGGAGCGGGTGGAGACAC 1739
Db 787 CAAGCCATTTCTGGGGCACTATCCACACTTCCACACTGAAGGTGTGGTGGGGGCAC 846
Qy 1740 TCACCTGATGTTATCACATGGCAGCGAGCTCAATATTTCTACCTCTCCACACCCCC 1799
Db 847 GCTCCGATGCTAATGATGCTGGCGAATTTAAATCTCCAGCTTCTACCAACCCC 906
Qy 1800 ACTATTCCTTATACCAATTAATAGGTTGCGAAGACACTTTAGACATGCTCATGACATGCCAC 1859
Db 907 ACCATTCTTTTCCCGTGATACAGAAGCCGACACATGGACATTTGATGGTGGCCAC 966
Qy 1860 CACTAGACAAACGATCCCGAGGATTTACAATTTTCTCAAGCCGATATCCGCCCGGC 1919
Db 967 CACTTGGATAAAAACATCAAGAAGATGTCCAGTTTGTCTGATTTCTAGGATTCGCCCCCAA 1026
Qy 1920 TCTATCGCGCTGAAGATGTCTCCATGATGCTGATGCTGATGCGATGACAGCTCGAT 1979
Db 1027 ACCATGCGCGCTGAGGACAAATCCACGATATGGGATTTTCTATACACAGCTCTGAC 1086
Qy 1980 TCCGACAGCATGGGGCTGCGAGCGAAGTATTCCTCGAACTTGGCAGACTGGCGGATAAG 2039
Db 1087 TCCGACAGCATGGGGCTGTAGCGAGGTATCATCCCGCACTTTGGCAACAGCGGACAA 1146
Qy 2040 AATAAAAAAGAAATTTGGTAGCTTCTGAAAGATGGCAAGATACGATAATTTCCGCAAT 2099
Db 1147 AACAAAAAAGAAATTTGGTCTGCTTGGCTGAGGAAAGGCGATAATGACAACTTCCGCATC 1206
Qy 2100 AAGGCTACATCTCCAATACATCTATCAACCCCGCTTTGACCCACCGCGTGGAGGATAT 2159
Db 1207 AAGGCTACATCTTCCAATACACCATCAACCCCGCTATTGGACACAGCGCATTTCTGAATAT 1266
Qy 2160 ATCGGCTCTGTGAAGAGGCAAGATCGCGACTTGGTGTGGGAATCTGCTTTTTT 2219
Db 1267 GTGGCTCTGTAGAAGTGGGCAAAATTCGCCGATTTGGTCTTGGAGTCTCGGCTCTTT 1326
Qy 2220 GCGGTAAAAACCCAAATTCGTGATCAAGGCGGTATGGTGTCTTCTGAAATGGGCGAT 2279
Db 1327 GGCATTAAACCCCAACATGATCATCAAGGCGGATTCATCGCACTTCTTCTCAATGGGCGAT 1386

Qy 2280 TCTAACGCGTCTGTGCCACTCCCAACCGGTTTATTACCCGAAATGTTTGGCATCAC 2339
Db 1387 GCCAATGCCCTTATCCCACTCCCAACCGGTGTTATTACCGGAAATGTTTGGCCACCAT 1446
Qy 2340 GGCAAGCGAAATTTGACACAGCATCACTTTTGTTCCTCAAGTCGCTATGAAAATGGC 2399
Db 1447 GGTAAAGCCAAATTTGACACCAATATCACTTTTGTATCCCAAGTGGCTTATGACAACGGC 1506
Qy 2400 GTCAAGAAAAAGCTGGGCTTTAGAGCGCAAGTTTCTACCGGTCAAAAACTGCCGTAACATC 2459
Db 1507 ATTAAAGAGAGTTGGCTTGCRAAGAGTGGTTTGGCAGTTAAAACTCGCGCAACATC 1566
Qy 2460 ACCAAGAAAGACTTCAAGTTCAACGACAAAAACGGCAAAAAATCACCGTCGATCCGAAACC 2519
Db 1567 ACCAAAAAGACCTCAAAATTCACGATGTTTACCGCACATCGAAGTCAATCTCTGAAACC 1626
Qy 2520 TTGAGGTCTTTGTAGATGCAAACTCTGCACCTCTTAAACCCACCTCGAAGTGCCTTA 2579
Db 1627 TACAAAGTTTAAAGTGGATGGCAAGAGTTTACTTCCAAAGCAGCGGATAAAATCAGCCTA 1686
Qy 2580 GCCCAGCGCTACACTTTTCTTCTAG 2603
Db 1687 GCACACTCTACAACTTGTCTTAG 1710

RESULT 15

ABR00816

ID ABR00816 standard; cDNA; 1719 BP.

XX ABA00816;

AC AC

XX 01-APR-2003 (first entry)

XX H. felis urease B coding sequence.

Gene; rat; IgG2a; light chain; E. coli; galactosidase; immunogen;
immune response; circulatory vessel; Gut Associated Lymphoid tissue;
GALT; immune response; antigen; Salmonella; Cholera; immunity;
Helicobacter pylori; HIV; Candida; P. gingivalis; gut; parasite; toxin;
hormone; hormone receptor; cancer; ss.

XX Helicobacter felis.

XX WO200296949-A1.

XX 05-DEC-2002.

XX 24-MAY-2002; 2002WO-AU000661.

XX 25-MAY-2001; 2001AU-00005241.

XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

XX McKenzie B, Boyle J, Lew A;

XX WPI; 2003-156789/15.

Raising an immune response in an animal by administering composition
comprising carrier and antigen bound to targeting moiety which binds to
receptor present in circulatory vessels in Gut Associated Lymphoid
tissue.

XX Disclosure; Page 27-28; 45pp; English.

This sequence encodes H. felis urease B. This sequence was used in the
method of the invention for raising an immune response in an animal. The
method comprises administering to the animal a composition comprising a
carrier and an antigen bound to a targeting moiety which binds to at
least one receptor present in circulatory vessels in Gut Associated
Lymphoid Tissue (GALT). The method is useful for raising an immune
response in an animal against antigens from Salmonella, Cholera,
Helicobacter pylori, HIV, Candida, P. gingivalis, gut parasites, gut

CC	associated toxins, gut hormones, gut hormone receptors or gut associated
CC	cancers. The method is useful for raising both mucosal and systemic
XX	immunity against any antigen used in the composition
SQ	Sequence 1719 BP; 490 A; 411 C; 421 G; 397 T; 0 U; 0 Other;
Query Match	31.3%; Score 902.4; DB 8; Length 1719;
Best Local Similarity	70.6%; Pred. No. 2e-192;
Matches 1203; Conservative	0; Mismatches 501; Indels 0; Gaps 0;
QY	900 AAAATGAAAAAACAAGATATGTAATACCTACGACCCACCAAGGCGATAAGTGGCG 959
DB	15 AAGATTTCACAAAGAAATATGTTCTATGATATGTTCCCACTACCGGGATCGTTAGA 74
QY	960 TTAGGAGATACCGATCTTTGGGAGAGTAGAATGACTATATACCACTATGGCGAAGAA 1019
DB	75 CTCGGGACACTGATTGATCTTTAGAGTGAGCATGATTGCACCACTTATGTGTAAGAG 134
QY	1020 CTTAAATTTGGCGGGTAAACTATCCGTTAGGGTATGGGTGAGCAATAGCCCTGAT 1079
DB	135 ATCAAAATTTGGGGCGGTAAACTATCCGTTAGGGTATGATCAAAACCAATAGCCCTAGC 194
QY	1080 GAAAAACCCCTAGATTAGTATCATCACTAACGCGATGATTATCGACTACACGGGATTTAC 1139
DB	195 TCTTATGAATTAGATTGGTGTCTCACTAACGCCCTCAATTGTGACTATACGGGCATTTAC 254
QY	1140 AAAGCCGACATTTGGATTAAAAACGGCAAAATCATGGCATTTGGCAAGGAGGAAACAAG 1199
DB	255 AAAGCCGACATTTGGATTAAAGACGCAAGATTTCAGGCATTTGGCAAGGAGGCAATAAG 314
QY	1200 GACATGAAGATGCGTAAAGCCCTCATATGGTGTGGGTGGGCAAGAAAGCACTAGCA 1259
DB	315 GACATGCAAGATGCGTAGATAATAATCTTTGGGTAGGTCTGTCTACAGAGGCTTTGGCA 374
QY	1260 GGGGAAGTATGATTATTCGCTGGGGATTCGATTTCACACCCCACTTCCTTTCTCCA 1319
DB	375 GCTGAGGCTTGATTGTAACCGTGTGGTATCGATACGCATATTCACATTTATCTCTCCC 434
QY	1320 CAACAAATTCCTACCGCTAGCAATGGCGTTTACAAACCATGTTTGGAGCGGCACAGGT 1379
DB	435 CAACAAATTCCTACTGCTTTTGGCAGCGGGTTACAAACCATGATTGGAGAGGCAAGGA 494
QY	1380 CTTGTAGATGGCAGAAATCGACTACTATCATCTCCGGGCAAAATGGAACTTGCACCGCATG 1439
DB	495 CTTGCGATGGCAGAAATGGACCACTCATCTCCGGACGCGTAAATCTTAAAGATG 554
QY	1440 TTGGCCGACGAGAGATTTCTATGAATGTGGGCTTTTGGCAAGGCAATAGCTCT 1499
DB	555 TTGGCTGACGCGAAGAAATACGCCATGAATCTAGGCTTTTGGCTAAGGGGAATGTGTCT 614
QY	1500 AGCAAAAAACAACCTTGTAGAACAGTAGAAGCGGGCGGATTTGTTTAAATTGCATGAA 1559
DB	615 TACGAACCTCTTTACCGGATCAGATTGAAGAGGGGGATTTGTTTAAATCCACGAA 674
QY	1560 GACTGGGGCAACAACAAGTCGATCGATCATCTGTTGAGCGTGGCAGATGAATACGAT 1619
DB	675 GACTGGGGAGCACACCTGCGAGCTATTTCACCACTGCTCAATGTCCGCGATGAATACGAT 734
QY	1620 GTCAAGTTTGTATCCACACCGATACAGTCAATGAGGAGGCTTATGTAGTACACCCCTA 1679
DB	735 GTCAAGTTGGCTATCCACACCGATACCGCTTTAAACGAGGGGGCTGTGTAGAGACACCCCTA 794
QY	1680 AATGCAATGAACGGCGGCCCATCCATCCCTACCACTTGAAGGGAGGGGTGGAGGACAC 1739
DB	795 GAGGCGATTTGGGGGGCGGCCATCCATCCATCTTCACACTGAAGGGGGCTGGGGGTGGACAC 854
QY	1740 TCACCTGATGTTATCAACATGGCAGGGAGCTCAATATTTCTACCCCTCCTCCACCAACCCCC 1799
DB	855 GCTCCAGATGTTATCAAAATGGCAGGGGAATTTAACTTTCTACCGCTCTACTTAAACCG 914
QY	1800 ACTATTCCTATACCAATTAATPACGGTTGCAAAACATTTAGACATGCTCATGACATGCCAC 1859
DB	915 ACCATTCTTTTCAACCAAAAACTGTAAGCCGAGCACATGGACATGTTTAAATGGTGTGCCAC 974

QY	1860 CACTAGACAAACGCAATCCGCGAGGATTTACAATTTTCTCAAGCCGATATCCGCCCCGGC 1919
DB	975 CACTTGGATAAAAGATATCAAGGAAGATGTGCAGTTTGCCTTTCGAGGATTCGCCCAA 1034
QY	1920 TCTATCGCGCTGAAGATGTCTCCATGATATCGGTGTGATCGGATGACAAGCTCCGAT 1979
DB	1035 ACTATCGCGCTGAAGACCAACTCCATGATACATGGGATCTTTCTATCACAGCTCCGAC 1094
QY	1980 TTCCAAGCAATGGGGGTGCGAGCGAAGTATTTCTCGAACTTTGGCAGACTGCGGATAAG 2039
DB	1095 TCTCAGCTATGGACGCGTAGCGAGGTGATCACACGCACTTTGGCAGACAGACAAA 1154
QY	2040 AATAAAAAAGAAATTTGGTAAGCTTCTGTAAGATGGCAAAAGATAACGATAATTTCCGAT 2099
DB	1155 AACAAAAAGAGTTTGGGGCTTTGAAAGAGGAAAAAGGCGATTAACGACAACTTCGATC 1214
QY	2100 AAGGGCTACATCTCCAATAACATATCAACCCCGCTTTGACCCACGCGTGAGCGAGTAT 2159
DB	1215 AAACGCTACATCTCTAAATACACATCAACCCCGCATCGGCATGGGATTTCTGACTAT 1274
QY	2160 ATCGGCTCTGTGAAGAGGCAAGATCGCCGACTTTGGTGTGTGGAATCTCTGCCCTTTT 2219
DB	1275 GTGGGCTCTGTGAAGTGGGCAATACGCGCACTCTGTGCTTTGGAGTCCGGCTTTCTTT 1334
QY	2220 GGCGTAAAAACCCAAATTCGTGATCAAGCGGTATGGTGTCTTCTCTGAAATGGGCGAT 2279
DB	1335 GGCAATTAAGCCCAATATGATTAATAAGGGCGGATTTATTGCGCTCTCTCAATGGGCGAT 1394
QY	2280 TCTAAACGCTCTGTGCCACTCCCAACCGGTTTATACCGCAAAATGTTTGGGCATCAC 2339
DB	1395 GCCAATGCGTCTATTTCCACCCCTCAGCCGCTTATTACCGTGAATGTTTGGACACCAT 1454
QY	2340 GGCAAGGCGAAATTTGACACGACATCACTTTTGTGTTTCCAAAGTCGCTATGAAAAATGGC 2399
DB	1455 GGGAAAAACAATTCGACCAATATCACTTTGTGTCCCAAGCGGCTTACAAGCGAGG 1514
QY	2400 GTGAAGAAAAAGCTTGGGCTTTAGAGCGCAAGTTCTACCGGTCAAAAACTGCGGTAAACATC 2459
DB	1515 ATCAAGAAAGAACTTAGGGCTTAGATCGGTGTATTTGCCAGTGAATAAACTGTGCAATATC 1574
QY	2460 ACCAAGAAAGACTTCAAGTTCAAGCAAAACGCAAAATCACCGTCGATCCGAAACCC 2519
DB	1575 ACTAAAAAGGACCTCAAAATTCAGCATGTGACCGCACATATTGATGTCAACCCCTGAAACC 1634
QY	2520 TTCGAGTCTTTGTAGATGCAAACTCTGCACCTCTAAACCCACCTCGCAAGTGCCTCTA 2579
DB	1635 TATAAGGTGAAGTGGATGGCAAGAGGTAACTCTTAAGCAGCAGATGAATTTAGCCTA 1694
QY	2580 GCCCAGCGCTACACTTTTCTCTAG 2603
DB	1695 GCGCAACTTTTATAATTTGTTCTAG 1718

Search completed: November 29, 2005, 00:34:36
Job time : 1579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 23:31:00 ; Search time 10210 Seconds
(without alignments)
13211.279 Million cell updates/sec

Title: US-09-904-994B-1
Perfect score: 2883
Sequence: 1 rrgagatttccarctt.....aaaaagtagaacacagg 2883

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156643

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_hic.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_gsal.*
10: gb_gsal2.*
11: gb_gsal3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	327.4	11.4	958	DR637438	DR637438 EST102806
2	324.2	11.2	946	DR637093	DR637093 EST102771
3	310.2	10.8	1509	CN50A8R8	BX816461 Arabidops
4	309	10.7	872	CO122599	CO122599 GR_Eb04F
5	306.2	10.6	925	DR633422	DR633422 EST102404
6	302.2	10.5	748	CX676965	CX676965 ydd29h04.
7	299.2	10.4	889	DR637386	DR637386 EST102801
8	296.4	10.3	796	DR633939	DR633939 EST102456
9	292	10.1	849	DR635599	DR635599 EST102622
10	279.6	9.7	766	CV195867	CV195867 CGF100344
11	274.2	9.5	986	CF826586	CF826586 EST703968
12	273	9.5	784	AZ935182	AZ935182 BJ_Ba000
13	262.6	9.1	990	CO027797	CO027797 EST806181
14	261.2	9.1	684	CF714948	CN907928 03109ABL
15	261	9.1	921	CF714948	CF714948 CCAF343TO
16	261	9.1	946	DN809412	DN809412 73954613
17	259.6	9.0	713	CK757406	CK757406 atr02-4ms
18	258.2	9.0	689	DR440006	DR440006 EST14909
19	255.2	8.9	781	DR916924	DR916924 EST110846
20	254.8	8.8	863	CD375065	CD375065 PTMM00226
21	254.4	8.8	945	DN808832	DN808832 76945303
22	249.4	8.7	757	BU027777	BU027777 QHG7107.Y

ALIGNMENTS

RESULT 1
DR637438

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

DR637438 958 bp mRNA linear EST 11-JUL-2005
EST1028063 FvM Gibberella moniliformis cDNA clone FVMAW64, mRNA
sequence.
DR637438.1 GI:70712272
Gibberella moniliformis
Gibberella moniliformis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
1 (bases 1 to 958)
Brown, D.W., Cheung, F., Proctor, R.H., Butchko, A.E., Zheng, L., Lee, Y.,
Utterback, T., Smith, S., Feldblum, T., Glenn, A.E., Plattner, R.D.,
Kendra, D.F., Town, C.D. and Whitelaw, C.A.
Analysis of 87,000 expressed sequence tags reveals alternatively
spliced introns in multiple genes of the fumonisin gene cluster
Unpublished (2005)
Contact: Brown, D.W.
USDA/ARS/NCAUR
USDA
1815 N. University St., Peoria, IL 61604, USA
Tel: 309 681 6230
Fax: 309 681 6689
Email: brown@ncaur.usda.gov
TIGR sequence name: FVMAW64TH
Seq primer: AAT TAA CCC TCA CTA AAG GG.

FEATURES
source

Location/Qualifiers
1..958
/organism="Gibberella moniliformis"
/mol_type="mRNA"
/strain="m3125"
/db_xref="taxon:117187"
/clone="FVMAW64"
/cissue_type="mycelia"
/clone_lib="FvM"
/note="Vector: pBlueScript II SK(+) XR; Site 1: EcoRI;
Site 2: XhoI; anamorph: Fusarium verticillioides. Library
FvM was prepared from pooled RNA obtained from a 48-hour
and a 72 hour, liquid GYAM culture from strain M-3125.
Cultures were vacuum filtered and the mycelial mats were
frozen in liquid nitrogen, ground to a powder, and then
added to Trizol Reagent (Invitrogen, Carlsbad CA) at
approximately 1 g mycelia per 10 ml Trizol. The cDNA was

CA264652 SCQGLB204
CN908212 030109ABL
CF708414 CCABR70TR
CA236757 SCEQFL505
CV902222 PD019F1.m
AG940016 Drosophil
DN810553 76954039
BG887260 EST513111
CF689903 CCADS40TF
CN910029 030124ABL
CA851837 D18B02.D1
CX557362 yda35h11.
AW618902 EST320888
CF690511 CCRAZ91TR
AZ935008 BJ_Ba000
BZ572656 msH2 2738
CF706674 CCAF8387O
AZ935371 BJ_Ba000
AG351733 131F03.Ma
AG939430 Drosophil
CF702527 CCAHQ83TF
AG941093 Drosophil

23 248.2 8.6 905 6 CA264652
24 247.8 8.6 661 7 CN908212
25 247.4 8.6 796 6 CF708414
26 240.8 8.4 789 8 CA236757
27 240.4 8.3 692 8 CV902222
28 239 8.3 643 10 AG940016
29 238.8 8.3 937 8 DN810553
30 237.2 8.2 634 2 BG887260
31 235.4 8.2 829 6 CF689903
32 234.2 8.1 626 7 CN910029
33 233.8 8.1 636 6 CA851837
34 226.2 7.8 708 8 CX557362
35 226 7.8 589 1 AW618902
36 225.2 7.8 747 6 CF690511
37 224.4 7.8 660 9 AZ935008
38 223.8 7.8 1275 9 BZ572656
39 222.4 7.7 890 6 CF706674
40 220.8 7.7 614 9 AZ935371
41 220 7.6 750 2 BG351733
42 219.2 7.6 555 1 AW618900
43 218 7.6 663 10 AG939430
44 218 7.6 878 6 CF702527
45 216.8 7.5 621 10 AG941093

ORIGIN	Query Match	Score	DB 8;	Length	958;
	Best Local Similarity	59.9%;	Pred. No. 3.5e-70;	Indels	2; Gaps
	Matches	565;	Conservative	0;	Mismatches 376;
QY	1209	GATGGCGTAAGCCCTCATATGTCGTGGGTGTGGGCACAGAAAGCACTAGCAGGGGAAGGT	1268		
Db	2	GACGGCGTAACAGAGGCGCATGTCGTGCGAAGCTGCACCGATGTCGTAGCAGGTGAAGGA	61		
QY	1269	ATGATTTATTACCGCTGGGGGAATCGATTACACACCCACTTCTCTTCTCCACAACTTC	1328		
Db	62	AAGATTTGTGACCGCGGGCGCTATCGACACGATATTCACTTTATCTGCCCTCAGCAAGTA	131		
QY	1329	CTTACCGCTCTAGCCAAATGGCGTTTACAACCATGTTTGGAGCGGCGCACAGCTCCTGTAGAT	1388		
Db	122	CCCGAAGCTCTTGCACTCTGGTGTAAACCACTATGCTTGGCGGTGATCCGSCCAATGCT	181		
QY	1389	GGCACGAATTCGCACTACTATCTCCGGGCAAAATGGAACATTGCACCGCATGTTTGGCGCA	1448		
Db	182	GGAAACGAACCACTACTTGTACGCCCTGGTGTCTCATTCATATGGCTCAAAATGTTGCAGCGC	241		
QY	1449	GCAGAGAGTATTCTATGAATGTTGGGCTTTTTCGGGCAGAAAGGCAATAGCTCTAGCAAAAA	1508		
Db	242	TGCGATCAGCTTCCATCAATATTGGTATTACTTGGCAAGGTAAATGATGTCTTCTGTAG	301		
QY	1509	CAACTTGTGAACAAGTAGAAGCGGCGCGGATTTGGTTTTAAATTTGCATGAAGACTCGGGC	1568		
Db	302	GGTCTGGCGATCAGGTCAATGCTGGTGTCTTGGCTCAAGCTTCATGAGGACTGGGGT	361		
QY	1569	ACAACCAAGTGGGATTCGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTT	1628		
Db	362	TGCACCTCTGCTGTTATGACGCTGTGCTCAGTGTCTGTGATGAATTCGATATTCAATGT	421		
QY	1629	TGTTATCCACACCATCAGTCAATGAGGCAAGTTATGTAGATGACACCCCTAAATGCAATG	1688		
Db	422	CTTATTACACATGACACGCTTAAACGAGTCTGGCTTTGTGCAATCTACGATGCGTGTCTTC	481		
QY	1689	AACGGCGCGCCCATCGATGCTTACCACATTTGAGGAGCGGCTGGAGACACTCACCTGAT	1748		
Db	482	AAGAACCGCACAAATTCATCTTATCACACAGAGGCTGCAGAGGTGGCCATGCTCCGAT	541		
QY	1749	GTTATCACCAATGGAGCGAGCTCAATTTCTACCTCTCCACACCCCCCACTAATTTCCC	1808		
Db	542	ATCATCTCCGTGTAGAGCATCAAAATGTTCTGCCATCATCGACCAACCCCTACAAGACCA	601		
QY	1809	TATACCAATTAACGGTTGCAGACACTTAGACATGCTCATGATGCCACCCACTAGAC	1868		
Db	602	TTCACACGCAATACTCTCGATGAGCATCTCGATATGCTTATGTTCTGCCATCACTTGTCC	661		
QY	1869	AAACGGATCCGCGAGGATTTACAATTTTCTCAAAGCCGATATCCGCCCGGCTCTATCGCG	1928		
Db	662	AAGNATATCCAGAGGATGTAGCCTTCGCGGAGAGCCGATTTCTGTGCTGAACCATTTGCT	721		
QY	1929	GCTGAAGATGTCTTCATGATATATGGTGTGATTCGCGATGACAAGCTCGGATTCGCAAGCA	1988		
Db	722	GCTGAGGATGTATTACACGACAAAGGGCGCTATCAGCATGATGAGCTCTGACTTCGAGGCT	781		
QY	1989	ATGGGCGTGCAGCGAAGTGATTCTTCGAACTTGGCAGACTCGGGATAGAAATAAAAA	2048		
Db	782	ATGGGCGTTGCGGAGAGGTGCTTTTAAAGAACATGGAATATCTGCGCATAAAAATAAGGTG	841		
QY	2049	GAATTTTGGTAACTTCTCGAAGATG - GCAAAGATAACGATAATTTCCGCATTAAGCGCT	2106		
Db	842	CAGAGGGTTGGTTTCGCGGAGGATGAAGGCACAGGGCTGATATGCGCGGTAAACGCT	901		
QY	2107	ACATCTCAAATACACTATCAACCCCGCTTTTGACCCACGGCGT	2149		
Db	902	ATGTCAGCAGTATACTATTAAACCCAGCTATTGCTCAGGGCTT	944		

RESULT 2	DR637093	EST1027718	FvM Gibberella moniliformis	946 bp	mRNA	linear	EST 11-JUL-2005
LOCUS	DR637093	EST1027718	FvM Gibberella moniliformis	946 bp	mRNA	linear	EST 11-JUL-2005
DEFINITION	DR637093	EST1027718	FvM Gibberella moniliformis	946 bp	mRNA	linear	EST 11-JUL-2005
ACCESSION	DR637093	EST1027718	FvM Gibberella moniliformis	946 bp	mRNA	linear	EST 11-JUL-2005
VERSION	DR637093.1	GI:70711927	FvM Gibberella moniliformis	946 bp	mRNA	linear	EST 11-JUL-2005
KEYWORDS	DR637093	EST1027718	FvM Gibberella moniliformis	946 bp	mRNA	linear	EST 11-JUL-2005
SOURCE	DR637093	EST1027718	FvM Gibberella moniliformis	946 bp	mRNA	linear	EST 11-JUL-2005
ORGANISM	DR637093	EST1027718	FvM Gibberella moniliformis	946 bp	mRNA	linear	EST 11-JUL-2005
REFERENCE	DR637093	EST1027718	FvM Gibberella moniliformis	946 bp	mRNA	linear	EST 11-JUL-2005
AUTHORS	DR637093	EST1027718	FvM Gibberella moniliformis	946 bp	mRNA	linear	EST 11-JUL-2005
TITLE	DR637093	EST1027718	FvM Gibberella moniliformis	946 bp	mRNA	linear	EST 11-JUL-2005
JOURNAL	DR637093	EST1027718	FvM Gibberella moniliformis	946 bp	mRNA	linear	EST 11-JUL-2005
COMMENT	DR637093	EST1027718	FvM Gibberella moniliformis	946 bp	mRNA	linear	EST 11-JUL-2005
FEATURES	DR637093	EST1027718	FvM Gibberella moniliformis	946 bp	mRNA	linear	EST 11-JUL-2005
source	DR637093	EST1027718	FvM Gibberella moniliformis	946 bp	mRNA	linear	EST 11-JUL-2005
ORIGIN	DR637093	EST1027718	FvM Gibberella moniliformis	946 bp	mRNA	linear	EST 11-JUL-2005

QY	1509	CAACTTGTAGAACAGTAGAGCGGCGCGATTGGTTTAAATTGCATGAAGACTGGGGC	1568
Db	302	GGTCTGCGCGATCAGGTCGAATGCTGGTCTTGGGCTCAAGCTTCAGGACTCGGGT	361
QY	1569	ACAACACCAAGTGCATCGATCACTCTTGAGCGTGGCAGATGAATACGATGTGCAAGTT	1628
Db	362	TGCATCTCCCTGCTATTGACGGTTGCTCAGTGTCTGTGATGAATCGATATCAATGT	421
QY	1629	TGTATCCACACCGCATACAGTCAATGAGGAGGTTATGTAGATGACACCCCTAAATGCAATG	1688
Db	422	CTTATTTCACACTGACACGGCTTAACGAGTCTGGCTTTGTGCAATCTACGATCGCTGCTTC	481
QY	1689	AAGGCGCGCCATCCATGCTTACCATGATGAGGAGCGGGTGGAGGACATCACTCGAT	1748
Db	482	AAGAACCGCACAAATTCATATTATCACACAGAGGGTGCAGAGGTGGCCATGCTCCGGAT	541
QY	1749	GTTATCACCATGCGAGCGAGTCAATATTTCTACCTCTCCACACACCCCACTATTCCC	1808
Db	542	ATCATCTCCGTGTAGCATCAAAATGTTCTGCCATCATCGACCAACCCCTACAAGACCA	601
QY	1809	TATACCATTAATACGGTTGCGAGAACACTTAGACATGCTCATGACATGCCACCACTTAGAC	1868
Db	602	TTTCACAGCAATACTCTCGATGAGCATCTCGATATGCTTATGGTCTGCCATCACTTGTCC	661
QY	1869	AAACGATCCGCGAGATTTACAATTTCTCAAAGCGTATCGCCCCGGCTCTATCGCG	1928
Db	662	AAGAATATCCAGAGGATGTAGCTTTCGGGAGAGCGGTATTCGTCTGAAACCACTTGT	721
QY	1929	GCTGAAGATGCTCCATGATATGGTGTGATCGCATGACAGCTCGATTCGATTCGCAAGCA	1988
Db	722	GCTGAGGATGTATTACACACAAAGCGCTATCAGCATGATGAGTCTGACTCGCAGGCT	781
QY	1989	ATGGGCGTGCAGGCGAAGTGATTCCTCGAACTTGGCAGACTCGGATGAAGATAAAAA	2048
Db	782	ATGGGCGGTGCGGAGAGTGTGTTTAAAGACATGGAATCTCGCATTAATAAATGATG	841
QY	2049	GAATTTGGTGAAGTCTCTCAAGAT--GGCAAGATAACGATAATTTCCGCATTAAGCGCT	2106
Db	842	CAGAGGGTGTGTCGGGAGGATGAGGCACAGGGGCTGATAATGCGGCTGTAAACGCT	901
QY	2107	ACATCTCAAATACACTATCAACCCCGCTTTGACCCACCGCGT	2149
Db	902	ATGTGACGAAATATCTATTACCCAGCTATTGCTCAGGGCTT	944
RESULT 3			
CNSOABR8			
LOCUS			
DEFINITION			
CNSOABR8 1509 bp mRNA linear HTC 06-FEB-2004			
Arabidopsis thaliana Full-length cDNA Complete sequence from clone			
GSLTPGH42D08 of Hormone Treated Callus of strain col-0 of			
Arabidopsis thaliana (thale cress).			
BX816461			
BX816461.1 GI:42471655			
HTC; GSLT cDNA.			
Arabidopsis thaliana (thale cress)			
Arabidopsis thaliana			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;			
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
1 (bases 1 to 1509)			
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,			
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,			
Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.			
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:			
A Combined Approach to Evaluate and Improve Arabidopsis Genome			
Annotation			
Unpublished			
2 (bases 1 to 1509)			
Genoscope.			
Direct Submission			
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :			
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr			
- Web : www.genoscope.cns.fr)			

COMMENT			
The sequences are based on single pass reads.			
Life Technologies (a division of Invitrogen) members carried out			
full-length libraries construction : Temple G.			
Genoscope members carried out sequencing and annotation : Castelli			
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,			
Schachter V., Weissenbach J., Salanoubat M.			
URGV INRA : Clepet C., Caboche M.			
Annotation is based on the June 2003 version of the Arabidopsis			
genome released by MIPS (Munich Information center for Protein			
Sequences). 5 prime and 3 prime are assembled with Phrap.			
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full			
length			
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.			
FEATURES			
source			
Location/Qualifiers			
1..1509			
/organism="Arabidopsis thaliana"			
/mol_type="mRNA"			
/db_xref="taxon:3702"			
/clone="GSLTPGH42D08"			
/tissue_type="Hormone Treated Callus"			
/ecotype="Col-0"			
/plasmid="pCMVSPORT_6"			
1..1509			
/gene="Atlg67550"			
ORIGIN			
Query Match 10.8%; Score 310.2; DB 4; Length 1509;			
Best Local Similarity 59.1%; Pred. No. 7.7e-66;			
Matches 602; Conservative 0; Mismatches 413; Indels 4; Gaps 4;			
QY	1587	GATCCTGCTTGAGCGTGCGAGATGAATACGATGTGCAAGTTTGTATCCACCGATACA	1646
Db	1	GACAATTGTTTGGCAGTTGCAGAAAGATATGATATCCAGGTGAACATTCTACTGCACC	60
QY	1647	GTCAATGAGCGAGTTATGTAGATGACACCTTAATGCAATGAACGGCGCCATCCAT	1706
Db	61	TTAAATGAATCCGGTTTGTGTGAGCACACTATCAATGCATTTCTGTGGGAGAACATACAT	120
QY	1707	GCTACCACTTGAAGGCGGTGGAGGACACTCACCTCATGTTATCACCATGGCAGGC	1766
Db	121	ACATATCAGATTAAGGTGCTGTGTGGACATGACACAGATATCATTAGATTGTGGA	180
QY	1767	GAGCTCAATATTCTACCTCTCTCCACACCCCACTATTCCCTATACATTAAATACGGTT	1826
Db	181	GTGAAAAATGTACTCCCGTCATCAACCAACCAACCGCCATATACGAAAAATCTGTA	240
QY	1827	GCAGAACCTTAGACATGTCTATGACATGCCACCACTAGACAAACGATCCGCGAGAT	1886
Db	241	GATGAACATCTTGACATGCTGTGTTGCCATCACCTTGACAAGAACATCCCCAGAAGAT	300
QY	1887	TTACAAATTTCTCAAAGCGGTATCCGCCCGGCTCTATCCGGCTGA-AGATGTGCTCCA	1945
Db	301	GTAGCTTTTGTGTAATCAAGAATAAGAGCTGAAACAATAGCTGCAGAGAGATATATGCA	360
QY	1946	TGATATGGTGTGATCGCATCAAGCTCGGATTTCGAAGCAATGGGGCGTCAGCGGA	2005
Db	361	TGATATGGGAGCAATCAGCATTTATCTCTCCGACTCACAAGCAATGGTTCGATTGGAGA	420
QY	2006	AGTGATCTCTCGAATCTGGCAGACTCGGGATAGATAAAAAAGAAATTTGGTAGCTTCC	2065
Db	421	GGTAATAAGCAAACTTGGCAGACAGCTGACAAGATGAAGGCTCAGAGGGGAGCCATTAA	480
QY	2066	TGAAGATGGCAAGATGAAGTAATTTCCGATTAAGC-GCTACATCTCCAAATACACTA	2124
Db	481	TCTTAACATGGCTGATGATTAACCAAGGATCAAGCATACATATTGCAAGTTACACTA	540
QY	2125	TCAACCCCGCTTTGACCCACGGCGTGAGCGAGTATATCGGCTCTGTGTGAAGAGGCGAAGA	2184
Db	541	TAAACCCAGCTATAGCCAATGGTTTGCACACTTAATTGGCTCGTTGAGGTGAAGAAGC	600
QY	2185	TGCGCGACTTGGTGGTGAATTCCTGCTCTTTT-TTGGCGTAAAAACCCAAATCGTGATC	2243
Db	601	TGGCTGATCTTGTATTATATGCGACCGCGCTTTCTTTCGGAGCAAAACCCAGAAATGATAATC	660

QY 2244 AAAGCGGTATGGTCTTCTCTGAAATGGCGGATTTAAACCGCTCTGTGCCACTCCC 2303
 |||||
 Db 661 AAAGGAGGCAATATAGCATGGCTTAATATGGCGGATGCAAAATGCGAGCATACCCACTCCT 720
 |||||
 QY 2304 CAACCGGTTTATACCGGGAATGTTTGGCATACGCGAAGCGGAAATTTGACACCAGC 2363
 |||||
 Db 721 GAGCCGGTGATATCAAGACCTTATGTTGGAGCATTTGGAAGGCGGAAAGTGAATACTCT 780
 |||||
 QY 2364 ATCACTTTTGTTCAAAGTCGCTTATGAAATGGCGTGAAGAAAGCTGGCTTAGAG 2423
 |||||
 Db 781 GTTGCAATTTGTCAGCAAGCTGCTTGAGAAAGGGGTAAAGAACTATATGGACTCAAG 840
 |||||
 QY 2424 CGCCAAGTTCTACCGTCAAAACTGCCGTATACATCACCAAGAAAGACTTCAAGTTCAAC 2483
 |||||
 Db 841 AAGAGGGTTGTAGCGGTGCCAAGCGTAGCGAGCTCACAAAATCTGCATGAAGCTGAAT 900
 |||||
 QY 2484 GACAAAACGGCAAAATACCGTGCATCGAATCCGAAACCTTCGAGGTCTTTGTAGATGCAAA 2543
 |||||
 Db 901 GATGCGCTTCCAGAGATCAACCGTGGATCCAGAGACATACGTTGTCAACGCAAAATGGCGAG 960
 |||||
 QY 2544 CT-CTGCACCTCTAAACCCACCTCGCAAGTGCCTCTAGCCAGCGCTACACTTTCTTCT 2601
 |||||
 Db 961 GTCTTAGCATGTGCCAGCGAATTCAGTCCCTCTCTCGCGCGGAATTTCTCTCT 1019
 |||||

RESULT 4
 LOCUS CO122599 872 bp mRNA linear EST 16-JUN-2004
 DEFINITION GR_Eb04F01.f GR_Eb Gossypium raimondii cDNA clone GR_Eb04F01 5',
 mRNA sequence.

ACCESSION CO122599
 VERSION CO122599.1 GI:48821286

KEYWORDS EST.
 SOURCE Gossypium raimondii;
 ORGANISM Gossypium raimondii

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eucosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 1 (bases 1 to 872)

AUTHORS Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
 Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
 Wing,R.A.

TITLE Global assembly of Cotton ESTs
 JOURNAL Unpublished (2004)
 COMMENT Contact: Rod A. Wing
 Arizona Genomics Institute
 The University of Arizona
 Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259

FEATURES
 source
 Location/Qualifiers
 1..872
 /organism="Gossypium raimondii"
 /mol_type="mRNA"
 /db_xref="taxon:29730"
 /clone="GR_Eb04F01"
 /tissue_type="floral"
 /dev_stage="3 to +3 DPA"
 /lab_host="DH10B"
 /clone_lib="GR_Eb"
 /note="Vector: pCMV.SPORT-6.1; Site_1: NotI; Site_2:
 EcoRV; Library made by invitrogen with RNA supplied by
 Wendle lab. Directional cloned into NotI-EV. Colonies
 plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Query Match 10.7%; Score 309; DB 7; Length 872;
 Best Local Similarity 59.74; Pred. No. 1.3e-65;
 Matches 519; Conservative 0; Mismatches 350; Indels 0; Gaps 0;

QY 1260 GGGGAAGGTATGATTTATACCGTGGGGGAATCGATTACACACCACCTTCCTTTCTCCA 1319
 |||||
 Db 2 GGAGAAGGATTTGATCGTAATCTGAGGGGCTATTGACTGTGTCATTTTATATGTCCT 61
 |||||
 QY 1320 CAACAAATTCCTTACCGCTCTAGCAATGGGTTTACAAACCATGTTTGGAGCGGCACAGGT 1379
 |||||
 Db 62 CAGTTGGTTCATGAGCTATTTCAAGTGGCATCACAACTATTGAGAGGTGGACCGGA 121
 |||||
 QY 1380 CTTGTAGATGGCAAGATCGACTACTATCACTCCGGGCAAAATGGAACTTTGCACCGCATG 1439
 |||||
 Db 122 CCAGCTGAAGGAACACGTGCAACTACTTTGACGCCAGCTCCGTGCGAAATGAAATGATG 181
 |||||
 QY 1440 TTGCGCGCAGCAGAAAGATTTCTATGAATGTGGCTTTTGGGCAAGGCAATAGCTCT 1499
 |||||
 Db 182 CTGCAGTCCACTGATGACTTCCCTCTAAATTTTGGCTTCACAGGGAAGGGAATGGTTCT 241
 |||||
 QY 1500 AGCAAAAAACAATTTGTAGAACAAAGTAGAAGCGGGCGGATTTGGTTTAAATTCGATGAA 1559
 |||||
 Db 242 AAACCTGAAGAAATACATGAAATTAATCAAAAGCCGAGCAATGGACTGAAACTGCATGAG 301
 |||||
 QY 1560 GACTGGGCGCAACACCAAGTGGCATCGATCACTGCTTGAGCGTGGCAGATGAATACGAT 1619
 |||||
 Db 302 GATTGGGGAACACCTGCTGCAATAGACAGTTGTTTGGCTGTTGCAGAACTATATGAT 361
 |||||
 QY 1620 GTCCAAAGTTTGTATCCACACCGATACAGTCAATGAGGACAGTTATGTAGATGACACCTA 1679
 |||||
 Db 362 ATCCAAGTTTAATATTCATCTACTGACACCTTGAACGAATCTGGAATTTGTGGAACAACAAT 421
 |||||
 QY 1680 AATGCAATGAACGGCGCGCATCCATGCTTACCACATTCAGGGAGCGGGTGGAGGACAC 1739
 |||||
 Db 422 AATGCAATTTAAAGGAAGAACTATTACACTTATCACAGTGAAGTGCCTGCTGGTGGTCT 481
 |||||
 QY 1740 TCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTCTTACCTCTCTCCACCAACCC 1799
 |||||
 Db 482 GCTCCAGATATCATCAAGTATGTGGCGTTAAAAACGCTCTCCCTTCGTCAACAAACCG 541
 |||||
 QY 1800 ACTATTCCCTATACCATTAATACGTTTGCAGAACACTTAGACATCTCATGACATGCCAC 1859
 |||||
 Db 542 ACTCGCCTTATATCTCCAATCTATATAGATGAACATCTTGACATGCTGATGTTGGCAT 601
 |||||
 QY 1860 CACTAGACAAACGCATCCGCGAGGATTTCAATTTTCTCAAGCCGCTATCCGCCCGGC 1919
 |||||
 Db 602 CACCTCAGCAAGGATATTCCTGAAGATGTAGCAATTCAGAAATTAGGCGCGGAA 661
 |||||
 QY 1920 TCTATCGCGCTGAAGATGTCTCCATGATATGGTGTGATCGCGATGACAAAGTCTCGAT 1979
 |||||
 Db 662 ACCATTGCTCCGAGACATATTCATGACACAGGGGCAATCAGCATTTATATCTTCAGAT 721
 |||||
 QY 1980 TCACAAGCAATGGGGCGTGCAGGCAAGTGAATTCCTCGAACTTGGCAGACTGCGGATAG 2039
 |||||
 Db 722 TCACAGCGGATGGGTGCGATTTGGCGAGTGATATGCAGAACTTGGCAAAACCGCCACAAG 781
 |||||
 QY 2040 AATAAAAAAGAAATTTGGTGAAGCTTCTGAAGATGGCAAGATACGATAATTTCCGCATT 2099
 |||||
 Db 782 ATGAATATCAAAAGGGGACCACTTGGTCTGCACTCAGCAACAGCAATCTAAGAAATC 841
 |||||
 QY 2100 AAGCGCTACATCTCCAATACACTATCA 2128
 |||||
 Db 842 AAAGCTTATATAGCAAAATACACCATATA 870
 |||||

RESULT 5

DR633422

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

DR633422 925 bp mRNA linear EST 11-JUL-2005

EST1024047 Fw Gibberella moniliformis cDNA clone FVMA534, mRNA

sequence.

DR633422

DR633422.1 GI:70708256

EST.

Gibberella moniliformis

Gibberella moniliformis

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.

REFERENCE 1 (bases 1 to 925)
Brown, D.W., Cheung, F., Proctor, R.H., Butchko, A.E., Zheng, L., Lee, Y.,
Utterback, P., Smith, S., Feldblyum, T., Glenn, A.E., Plattner, R.D.,
Kendra, D.F., Town, C.D. and Whitelaw, C.A.

TITLE Analysis of 87,000 expressed sequence tags reveals alternatively
spliced introns in multiple genes of the fumonisin gene cluster

JOURNAL Unpublished (2005)
COMMENT Contact: Brown, D.W.
USDA/ARS/NCAUR

1815 N. University St, Peoria, IL 61604, USA

Tel: 309 681 6230

Fax: 309 681 6689

Email: brown@ncaur.usda.gov

TIGR sequence name: FVMA534TH

Seq primer: AAT TAA CCC TCA AAG GG.

Location/Qualifiers

1..925

/organism="Gibberella moniliformis"

/mol_type="mRNA"

/strain="m3125"

/db_xref="taxon:117187"

/clone="FVMA534"

/tissue_type="mycelia"

/note="Vector: pBlueScript II SK(+); Site 1: EcoRI;
Site 2: XhoI; anamorph: Fusarium verticillioides. Library
FvM was prepared from pooled RNA obtained from a 48-hour
and a 72 hour, liquid GYM culture from strain M-3125.
Cultures were vacuum filtered and the mycelial mats were
frozen in liquid nitrogen, ground to a powder, and then
added to Trizol Reagent (Invitrogen, Carlsbad CA) at
approximately 1 g mycelia per 10 ml Trizol. The cDNA was
directionally ligated into the pBlueScript II SK(+) XR
vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

vector (cDNA Synthesis Kit; Stratagene)."

QY 1792 CCACCCCACTATTCCCTATACCAATTAATACGGTTGCAGAACACTTACAGATGCTCATGA 1851
Db 490 CCAACCTTACAAGACCAATTCACACGCAATCTCTCGATGAGCATCTCGATATGCTTATGG 549
QY 1852 CATGCCACCACTAGACAAACGATCCGCGAGGATTTTACAAATTTTCTCAAAGCCGTATCC 1911
Db 550 TCTGCCATCACTTGTCCAGAAATATCCAGAGGATGTAGCCTTCGCGGAGAGCCGTATTC 609
QY 1912 GCCCCGGCTCTATCGCGGCTGAAGATGTCTCCATGATATGGGTGTGATCGCGATGACAA 1971
Db 610 GTGCTGAACCAATTTGCTGTGAGGATGTATTACAGCAAAAGGCGCTATCAGCATGATGA 669
QY 1972 GCTCGGATTCGCAAGCAATGGGCGGTGCAGCGAAGTATTCCTCGAACTTTGGCAGACTG 2031
Db 670 GCTCTGACTCGCAGGCTATGGGCGGTTTTCGCGAGAGGTCGTTTTTAAGAACATGGAATCTG 729
QY 2032 CGGATGAAGATAAAAAAGATTTGGTAAGCTTCTCTGAAGAT--GGCAAGATAACGATA 2088
Db 730 CGCATAAATAAAGGTGCAGAGGGTTGGTTGCGGAGGATGAGGACACAGGGGCTGATA 789
QY 2089 ATTTCGCGCATTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACGCG 2148
Db 790 ATGCGGCTGTAAACGCTATGTGAGCAAGTATCTATTACCCAGCTATTCTCAGGGCT 849
QY 2149 TGAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGATCGCGGA-CTTGGTGGTGTGGAAT 2207
Db 850 TTGGACATGTCTATGGGAGCATTTGAGGTTGGAAGTTTGTGATCTTGTCTTTTGGGAT 909
QY 2208 CTGCGCTTTTGGC 2222
Db 910 CTGCGTGGTGGC 924

RESULT 6
CX676965
LOCUS
DEFINITION
yddd29h04.y1 Sea urchin EST Lib1 Strongylocentrotus purpuratus cDNA
clone ydd29h04 5' similar to TR:Q9SR52 Q9SR52 PUTATIVE UREASE. ;,
mRNA sequence.

ACCESSION CX676965
VERSION CX676965.1 GI:57936920
KEYWORDS EST.
SOURCE Strongylocentrotus purpuratus
ORGANISM Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoidea;
Strongylocentrotidae; Strongylocentrotus.

REFERENCE 1 (bases 1 to 748)
AUTHORS Coffman, J.A., Robertson, A.J., Clifton, S., Pape, D., Hillier, L.,
Martin, J., Wylie, T., Dante, M., Meyer, R., Theising, B., Bowers, Y.,
Gibbons, M., Ronko, I., Teagareishvili, R., Ritter, B., Kennedy, S. and
Wilson, R.

TITLE Washu Sea Urchin EST Project
JOURNAL Unpublished (2004)
COMMENT Contact: Dr. James A. Coffman
Washu Sea Urchin EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
DNA sequencing by: Washington University Genome Sequencing Center
Seq primer: -28RppOT
High quality sequence stop: 680.

Location/Qualifiers
1..748
/organism="Strongylocentrotus purpuratus"
/mol_type="mRNA"
/db_xref="taxon:7668"
/clone="ydd29h04"
/lab_host="DH10B"
/clone_lib="Sea urchin EST Lib1"
/note="Vector: pCMVSPORT6.1 (Invitrogen); Site_1: Not1;

FEATURES
source
1..748
/organism="Strongylocentrotus purpuratus"
/mol_type="mRNA"
/db_xref="taxon:7668"
/clone="ydd29h04"
/lab_host="DH10B"
/clone_lib="Sea urchin EST Lib1"
/note="Vector: pCMVSPORT6.1 (Invitrogen); Site_1: Not1;

Site_2: Smal; Arrayed normalized library of full-length cDNAs representing blastula stage transcriptome of the sea urchin Strongylocentrotus purpuratus, cloned into the vector pCMVSPORT6.1 (Invitrogen)"

ORIGIN		Query Match	10.5%;	Score 302.2;	DB 8;	Length 748;			
		Best Local Similarity	63.1%;	Pred. No. 6.3e-64;					
		Matches 466;	Conservative 0;	Mismatches 273;	Indels 0;	Gaps 0;			
QY	1648	TCAATGAGCGAGTATATAGATGACACCCCTAAATCAATGAACGCGCGCCATCATG	1707						
DB	9	TCATGAATCTCTCTGTGTGAAGATACCATAGCTCTTTGAGGCGCGACCATCCACA	68						
QY	1708	CCTACACATTTGAGGAGCGGGTGGAGGACACCTCACCTGATGTATTACCATGCGCAGCG	1767						
DB	69	CCTATCACTCGAGGGTGCAGGAGGAGGCCACGCCCTGACATCATGAAGGTAGTGGCG	128						
QY	1768	AGTCAATATTCTACCTCTCTCCACACCCCACTATTTCCTATACCATTAATACGGTTG	1827						
DB	129	TGCCCAATGTATTACCTCTGTCTACCAACCCACCCGCCCTTTCACAGTGAACACCATCG	188						
QY	1828	CAGAACATTTAGACATGCTCATGACATGCCACCACTAGACAAACGATCCCGGAGATT	1887						
DB	189	ATGAACACCTGGATATGCTCATGGTCTGCCACCATCTTGACAAGATCTTAAGAGAGACG	248						
QY	1888	TACAATTTTCTCAAGCCGCTATCGGCCCGCTCTATCGCGGCTGAAGATGTCTCCATG	1947						
DB	249	TTGCCCTTGCAGATCTCGCATCAGAGCGGAACCATTTGCTGTGAAGATATCTTTCATG	308						
QY	1948	ATATGGGTGTGATCGGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGGGCAAG	2007						
DB	309	ATCTGGGAGCAATCAGTATTGTGGCTCTCCGATTCGCAGGCGCATGGGGCGTGTGGTGAGG	368						
QY	2008	TGATTTCTCGAATCTGCGACATCGGATGAAGATAAAGAAATTTGGTAAAGCTTCTG	2067						
DB	369	TGATCACCCTGTACATGGCAGCGCAGATGAAGATGAAGATTTTCAGAGGCGAGCTGTGAG	428						
QY	2068	AGATGGCAAGATGAACGATAATTTTCGCATTAAGCGCTATCATCTCCAAATACACTATCA	2127						
DB	429	AGAAACGGGTGACACGACAACTTAAGGTCGACGTTTACATTTGCCAGTACAGGTCA	488						
QY	2128	ACCCCGCTTTGACCCACCGCGTGAAGGATATATCGGCTCTGTGGAGAGGGCAAGATCG	2187						
DB	489	ACCCGGCATTGCCCCACCGCATGGGCCATCTCTATTGGCTCTGTACAGTTGGTAAATGG	548						
QY	2188	CCGACTTGGTGTGGAATCTCGCTCTTTTGGCGTAAACCCAAATCGTGATCAAG	2247						
DB	549	CTGACCTGGTTCTCTGGAATCCGGCATTTCTCGAGCTAAACCTGACCTCATCATCAAG	608						
QY	2248	GCGETATGGTGTCTCTGMAATGGCGGATTTCTAAGCGGTCTGTGCCCACTCCCCAAC	2307						
DB	609	GAGGTTACATTTGTTGGGCACAAATGGGCATGCGCAATGATCATCCGACCCCTGAC	668						
QY	2308	CGGTTTATTACCGCGAAATGTTTGGGATCAGCGCAAGGGGAAATTTGACACCCAGCATCA	2367						
DB	669	CTGTCAAGACAGGAAATGTTTGGTCTTACGGGAAGTCCATCGGTGAGAACTCTGTCA	728						
QY	2368	CTTTTGTTCGAAGTCGC	2386						
DB	729	TCTTTGTGTCGAAGCAGC	747						

RESULT 7
DR637386
LOCUS
DEFINITION
EST1028011 FvM Gibberella moniliformis cDNA clone FVMAV95, mRNA
889 bp mRNA linear EST 11-JUL-2005
ACCESSION
DR637386
VERSION
DR637386.1 GI:70712220
KEYWORDS
EST.
Gibberella moniliformis
SOURCE
Gibberella moniliformis

REFERENCE
AUTHORS

TITLE

JOURNAL
COMMENT

FEATURES
source

1. 889
/organism="Gibberella moniliformis"
/mol_type="mRNA"
/strain="m3125"
/db_xref="taxon:117187"
/clone="FVMAV95"
/tissue_type="mycelia"
/clone_lib="FvM"
/notes="vector: pBluescript II SK(+) XR; Site 1: EcoRI;
Site 2: XhoI; anamorph: Fusarium verticillioides. Library
FvM was prepared from pooled RNA obtained from a 48-hour
and a 72 hour, liquid GYAM culture from strain M-3125.
Cultures were vacuum filtered and the mycelial mats were
frozen in liquid nitrogen, ground to a powder, and then
added to TRIzol Reagent (Invitrogen, Carlsbad, CA) at
approximately 1 g mycelia per 10 ml TRIzol. The cDNA was
directionally ligated into the pBluescript II SK(+) XR
vector (cDNA Synthesis Kit; Stratagene)."

ORIGIN

		Query Match	10.4%;	Score 299.2;	DB 8;	Length 889;			
		Best Local Similarity	59.5%;	Pred. No. 3.7e-63;					
		Matches 524;	Conservative 0;	Mismatches 353;	Indels 3;	Gaps 1;			
QY	1312	TTTCTCCACAAACAATCCCTACCGCTCTAGCCAATGGGTTTACACCAATGTTTGAGCGG	1371						
DB	10	TCTGCCCTCAGCAAGTACCCGAGCTCTTGATCTGGTGTAAACCACTGCTTGGCGGTG	69						
QY	1372	GCACAGTCTCTGTATAGGCGACGAATCGGATCTACTATCACTCCGGGCAATGGACATTGC	1431						
DB	70	GTACCGGCCCAAGTGTGTGAACGAAACCACTTGTACGCTGTGTCTCATTTACATGC	129						
QY	1432	ACGCGATGTTGCGCGCAGCAAGAGTATTCTATGAATGTGGGCTTTTGGGCAAGGCA	1491						
DB	130	GTCAAAATTTGCGAGCGGTGCGATCAGCTTCCTATCAATATTGGTATTACTGGCAAGGTA	189						
QY	1492	ATAGCTCTAGCAAAAAACAATTTGTAGAAAGTAGAAGCGGGCGGATTTGGTTTAAAT	1551						
DB	190	ATGATAGTTCTCTGAGGGTCTGCGGATCAGGTCAATGCTGTGCTTGTGGCCTCAAGC	249						
QY	1552	TGATGAAGACTGGGCGACAAACCAAGTCGATCGATCACTGCTTGAGCGTGGCAGATG	1611						
DB	250	TTCATGAGGACTGGGGTTGCACTCCTGCTGTATTGACGCTTGTCTAGTGTCTGTGATG	309						
QY	1612	AATACGATGTCAAGTTTGTATTCCACCGGATACAGTCAATGAGCGAGGTATTGTAGATG	1671						
DB	310	AATTCGATATTCAATGTCTTATTACATGACACGCTTAAACGATCTGCTTTGTGCAAT	369						
QY	1672	ACACCTCAAAATGCAATGAACCGGGCGGCATCCATGCCCTACACATTTGAGGGAGCGGTG	1731						
DB	370	CTACGATCGTGTCTTCAAGAACCGCAATTCATACTTATCACACAGAGGGTGCAGGAG	429						
QY	1732	GAGGACACTCACTGATGTTTATCACCATGGCAGGCGAGCTCAATATTCTACCTCTCTCA	1791						

Db 430 GTGGCCATGCTCGGATATCATCTCGTGGTAGAGCATCAAAATGTTCTCCCATCATCGA 489
QY 1792 CCACCCCACTATTCCCTATACCAATTAATACCGTTGCGAGAACACTTAGACATGCTCATGA 1851
Db 490 CCAACCCCTACAAGACCAATTCACACGCAATCTCTCGATGAGCATCTCGATATGCTTATGG 549
QY 1852 CATGCCACACCTAGACAAACGATCGCGGAGGATTACAAATTTCTCAAAGCGGTATCC 1911
Db 550 TCTGCCATCACTTGTCTCAAGATATATCCAGAGGATGTAGCTTTCGGGAGAGCGGTATTC 609
QY 1912 GCCCGCGTCTATCGCGGTGAAGATGTCTCCATGATATGGGTGTGATCGCGATGACAA 1971
Db 610 GTGCTGAACCAATGTCTGTGAGATGTATTACGACAAAGCGCTATCAGCATGATGA 669
QY 1972 GCTCGGATTTCGAAGCAATGGGCGGTGCGAGGCAAGTGAATTCCTCGAACTTTGCAGACTG 2031
Db 670 GCTCTGACTCGCAGGCTATGGGCGGTTCGGAGAGGTTCGTTTAAAGACATGGAATACTG 729
QY 2032 CGGATGAAGATTAAGAAATTTGGTAAGCTTCTCTGAAGAT---GGCAAGATAACGATA 2088
Db 730 CGCATAAATAAGGTGCGAGGGGTGGTTCGGGAGGATGAGGCGACAGGGGCTGATA 789
QY 2089 ATTTCGGCATTAAGCGCTACATCTCCAAATACACTATCAACCGCTTTGACCCAGCGG 2148
Db 790 ATGCGGTGTAAACAGCTATGTGACGAAGTACTATTAAACCCAGCTATTGCTCAGGGCT 849
QY 2149 TGAGCGAGTATATCGCTCTGTGGAAGAGGCGCAAGATCGC 2188
Db 850 TTGGACATGTCATTGGAAGCATTGAGTTGGAAGTTTGC 889

RESULT 8
LOCUS DR633939 796 bp mRNA linear EST 11-JUL-2005
DEFINITION EST1024564 FvM Gibberella moniliformis cDNA clone FvMA846, mRNA
sequence.
ACCESSION DR633939
VERSION DR633939.1 GI:70708773
KEYWORDS EST.
SOURCE Gibberella moniliformis
ORGANISM Gibberella moniliformis
REFERENCE 1 (bases 1 to 796)
AUTHORS Brown,D.W., Cheung,F., Proctor,R.H., Butchko,A.E., Zheng,L., Lee,Y.,
Utterback,T., Smith,S., Feldblyum,T., Glenn,A.E., Flatner,R.D.,
Kendra,D.F., Town,C.D. and Whitelaw,C.A.
TITLE Analysis of 87,000 expressed sequence tags reveals alternatively
spliced introns in multiple genes of the fumonisin gene cluster
JOURNAL Unpublished (2005)
COMMENT Contact: Brown, D.W.
USDA/ARS/NCAR
USDA 1815 N. University St, Peoria, IL 61604, USA
Tel: 309 681 6230
Fax: 309 681 6689
Email: brown@ncar.usda.gov
TIGR sequence name: FvMA846TH
Seq primer: AAT TAA CCC TCA AAG GG.
Location/Qualifiers
1..796
/organism="Gibberella moniliformis"
/mol_type="mRNA"
/strain="m3125"
/db_xref="taxon:117187"
/clone="FvMA846"
/tissue_type="mycelia"
/clone_lib="FvM"
/note="Vector: pBlueScript II SK(+) XR; Site_1: EcoRI;
Site_2: XhoI; anamorph: Fusarium verticillioides. Library
FvM was prepared from pooled RNA obtained from a 48-hour
and a 72 hour, liquid GYAM culture from strain M-3125.
Cultures were vacuum filtered and the mycelial mats were

ORIGIN
Query Match 10.3%; Score 296.4; DB 8; Length 796;
Best Local Similarity 60.8%; Pred. No. 1.8e-62;
Matches 483; Conservative 0; Mismatches 311; Indels 0; Gaps 0;
QY 1243 GCACAGACCACTAGCAGGGAAGGTATGATTATACCGTGGGGATCGATTACACACA 1302
Db 3 GCACGAGTGTGAGCAGGTGAAGAAAGATTGACCGCGGCGCTATGACACGCATA 62
QY 1303 CCACATTCCTTTCTCCACAAATTCCTCTACCGCTCTAGCCAAATGGCGGTACAAACCATGT 1362
Db 63 TTCATTTATCTGCCCTCAGCAAGTACCCGAGCTCTTGCACTGGTGTACCAACCATGC 122
QY 1363 TTGGAGCGGCACAGGTCTCTGTAGATGGCAGCAATGCGACTACTATCACTCCGGCAAT 1422
Db 123 TTGGCGTGTGTACCGGCCCAAGTGTGGAACGAACCACTACTTGTACGCTGTGTCTC 182
QY 1423 GGAACCTTGACCGCATGTTCGCGCAGCAGAGATATTCTATGAATGGGCTTTTGG 1482
Db 183 ATTACATGCGTCAAAATGTTGCGAGCGTGCATCAGTTCCTTATCAATATTGGTATCTG 242
QY 1483 GCAAGCGCAATAGCTCTTAGCAAAAACAACTTGTAGAACAGTAGAGCGGCGGATTTG 1542
Db 243 GCAAGCGTAATGATAGTTCCTCGAGGTCTGCCGATCAGGTCAATGCTGTGTCTGTG 302
QY 1543 GTTTTAAATTCATGAAGACTGGGGCACAAACACCAAGTGCATCGATCACTGCTTGAGCG 1602
Db 303 GCCTCAAGCTTCATGAGGACTGGGTTGCACTCTCTGCTGCTATTGACGCTTGTCTCAGTG 362
QY 1603 TGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGAGGTT 1662
Db 363 TCTGTGATGAATTCGATATTCAATGCTTATTCACACTGACACGCTTAAACGAGTCTGGCT 422
QY 1663 ATGTAGATGACACCTTAATGCAATGACCGGGCGCATCCATCGCTTACCACTTGAGG 1722
Db 423 TTGTCGAATCTACGATCGCTGCTTTCAAGAACCGCAAAATTCATACTTATCACACAGAGG 482
QY 1723 GACGGGTGGAGGACACTCACCTGATGTTATCACTATGCGAGGCGAGCTCAATATTCTAC 1782
Db 483 GTGCAGAGGTGGCATGTCTCCGATATCATCTCCGTGGTAGACATCAAAATGTTCTGCG 542
QY 1783 CTTCTCCACCCACCCCACTATTTCCTTATACCAATTAATACGGTTGCGAGAACACTTAGACA 1842
Db 543 CATCATCGACCAACCCCTACAAGACCAATTCACACGAATACTCTCGATGAGCATCTCGATA 602
QY 1843 TGCTCATGATCGCACCACTACACCAAGCAACGATCCCGGAGGATTACAAATTTCTCAA 1902
Db 603 TGCTTATGTTGCTGCCATCACTTGTCTCAAGAAATATCCAGAGGATGTAGCCTTCGCGGAGA 662
QY 1903 GCGTATCCGCGCGCTCTATCGCGCTCAAGATGTCTCCATGATATGATGATGATGATCG 1962
Db 663 GCGTATTCGTCTGAAACCAATTCGTCTGAGATGATTATACACGACAAAGCGCGTATCA 722
QY 1963 CGATGACAAAGCTCGGATTTCGCAAGCAATGGGGCGTGCAGGCGAAGTATTCTCTCGAACTT 2022
Db 723 GCATGATGAGTCTGACTCGAGGCTATGCGCGCTATGCGGAGAGTCTGTTTAAAGAACAT 782
QY 2023 GGCAGACTCGGAT 2036
Db 783 GGAATACTCGCAT 796

RESULT 9
LOCUS DR635599 849 bp mRNA linear EST 11-JUL-2005
DEFINITION EST1026224 FvM Gibberella moniliformis cDNA clone FvMAJ96, mRNA
sequence.

ACCESSION DR635599
 VERSION DR635599.1 GI:70710433
 KEYWORDS EST.
 SOURCE Gibberella moniliformis
 ORGANISM Gibberella moniliformis
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
 REFERENCE 1 (bases 1 to 849)
 AUTHORS Brown, D.W., Cheung, F., Proctor, R.H., Butchko, A.E., Zhang, L., Lee, Y., Utterback, T., Smith, S., Feldblyum, T., Glenn, A.E., Plattner, R.D., Kendra, D.F., Town, C.D. and Whitelaw, C.A.
 TITLE Analysis of 87,000 expressed sequence tags reveals alternatively spliced introns in multiple genes of the fumonisin gene cluster
 JOURNAL Unpublished (2005)
 COMMENT Contact: Brown, D.W.
 USDA/ARS/NCAR
 1815 N. University St, Peoria, IL 61604, USA
 Tel: 309 681 6230
 Fax: 309 681 6689
 Email: brown@ncar.usda.gov
 TIGR sequence name: FVMAJ96TH
 Seq primer: AAT TAA CCC TCA CTA AAG GG.
 FEATURES
 source Location/Qualifiers
 1..849
 /organism="Gibberella moniliformis"
 /mol_type="mRNA"
 /strain="m3125"
 /db_xref="taxon:117187"
 /clone="FVMAJ96"
 /tissue_type="mycelia"
 /clone_lib="FVM"
 /note="Vector: pBluescript II SK(+); Site 1: EcoRI; Site 2: XhoI; anamorph: Fusarium verticillioides. Library FVM was prepared from pooled RNA obtained from a 48-hour and a 72 hour, liquid GYM culture from strain M-3125. Cultures were vacuum filtered and the mycelial mats were frozen in liquid nitrogen, ground to a powder, and then added to Trizol Reagent (Invitrogen, Carlsbad CA) at approximately 1 g mycelia per 10 ml Trizol. The cDNA was directionally ligated into the pBluescript II SK(+) XR vector (cDNA Synthesis Kit; Stratagene)."
 ORIGIN
 Query Match 10.1%; Score 292; DB 8; Length 849;
 Best Local Similarity 60.2%; Pred. No. 2.3e-61;
 Matches 503; Conservative 0; Mismatches 330; Indels 3; Gaps 1;
 QY 1317 CCACAAATTCCTACCGCTCTAGCCCAATGGCGTTACAAACCATGTTTGGAGCGGCAC 1376
 DB 3 CCTCAGCAAGTACCGAAGCTCTTGCACTCTGGTCTAACCAACCATGCTGGCGGTGATAC 62
 QY 1377 GGTCTGTAGATGCGCAATGCGACTACTATCACTCGGGCAATGGAACTTGCACCGC 1436
 DB 63 GGCCCCAAGTCTGGAACGACGCAACTACTTGTACGCTGTGTCTAATCATGCGTCAA 122
 QY 1437 ATGTTGGCGCAGCAGAAAGATTTCTATGAATCTGGCTTTTGGCAAGGCAATAGC 1496
 DB 123 ATGTTGCGCGGTGCGATCAGCTTCCTATCAATTTGGTATCTAGCAAGGTAAATGAT 182
 QY 1497 TCTAGCAAAAAACAACTTTGTAGAACAGTGAAGCGGCGGATGTTGTTTAAATTCAT 1556
 DB 183 AGTTCTCTCAGGGTCTGCGGATCAGTCAATGCTGGTGTGTGGCTCAAGCTTTCAT 242
 QY 1557 GAAGACTGGGCAACACCAAGTGGCATCGATCACTGCTGAGCGTGGCGAGATATAC 1616
 DB 243 GAGGACTGGGGTTGCACTCCCTGCTATTGACGCTTGTCTCACTGCTGTGTGAATTC 302
 QY 1617 GATGTGAAGTTTCTATCCACCGCATACATCAATGAGGCGAGGTATGTAGATGACACC 1676
 DB 303 GATATTCATATGCTTTATTCACATCTGACACGCTTAACAGTCTGCTTTGTGCAATCTACG 362
 QY 1677 CTTAATGCAATGAACGGGCGCGCCATCCATGCCCTACCACTTTGAGGGAGCGGTGGAGGA 1736

Db 363 ATCGCTGCTTTCAAGAACCGCAATTCATCTTATCACACAGAGGTGCGAGAGTGGC 422
 QY 1737 CACTCACCTGATGTTATCACCATGGCAGGAGCTCAATATTCTACCTCTCCACCAACC 1796
 Db 423 CATGCTCCGATATCATCTCCGTGGTAGAGCATCAAAATGTTCTGCCATCATCGACCAAC 482
 QY 1797 CCACACTATCCCTATACCATTAATACGGTTGCAGAACACTTAGACATGCTCATGACATGC 1856
 Db 483 CTTACAAGACCATTTCAACGCAATACTCTCGATGAGCATCTCGATATGCTTATGGTCTGC 542
 QY 1857 CACCACCTAGACAAACGCATCCCGAGGATTTCAAAATTTCTCAAAGCCGCTATCCGCCCC 1916
 Db 543 CATCATTTGTTCAAGATATATCCAGAGATGACCTTCCGGAGAGCCGTATTGCTGCT 602
 QY 1917 GGCTCTATCGCGCTGAAGATGTCTCCATGATATGGTGTGATCGGATGACAAGCTCG 1976
 Db 603 GAAACCATTTGCTGTGAGGATGATTACACGACAAAGGCGCTATCAGCATGATGAGCTCT 662
 QY 1977 GATTCCGACCAATGGGCGTGCAGGCAAGTGAATCTCGAACTTGCAGACTGCGGAT 2036
 Db 663 GACTCGCAGCTATGGGCCGTTCGGAGAGGTGCTTTTAAGAACATGGAATACTGCGCAT 722
 QY 2037 AAGAATAAAAAAGAAATTTGTTAAGCTTCTCTGAAGAT---GGCAAAGATAACGATAATTTC 2093
 Db 723 AAAATAAGTGTGACAGGGGTGTTGCCGAGAGATGAGGSCACAGGGGCTGATATCGG 782
 QY 2094 CGCATTAAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTTGACCCACGGCGT 2149
 Db 783 CGTGTAACGCTATGTGACCAAGTATACTATTAAACCAGCTATTGTCTCAGGGCTT 838
 RESULT 10
 CV195867 766 bp mRNA linear EST 15-SEP-2004
 LOCUS CGF1003441 All Seed coat from mid-season walnut embryos collected
 DEFINITION Aug 1 Juglans regia cDNA clone WSC0002_IVF_A11 5', mRNA sequence.
 ACCESSION CV195867
 VERSION CV195867.1 GI:52124704
 KEYWORDS EST.
 SOURCE Juglans regia (English walnut)
 ORGANISM Juglans regia
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fagales; Juglandaceae; Juglans.
 REFERENCE 1 (bases 1 to 766)
 AUTHORS Muir, R., Baek, J., Leslie, A., Cook, D. and Dandekar, A.
 TITLE Analysis of genes expressed in walnut seed coat tissue
 JOURNAL Unpublished (2004)
 COMMENT Contact: Abhaya Dandekar, PhD
 CAES Genome Facility
 UC Davis, Department of Pomology
 One Shields Ave, Davis, CA 95616, USA
 Tel: 530 752 7784
 Fax: 530 752 8502
 Email: andandekar@ucdavis.edu
 Seq primer: WSCF-TCCGAGATCTGGACGAGC.
 FEATURES
 source Location/Qualifiers
 1..766
 /organism="Juglans regia"
 /mol_type="mRNA"
 /cultivar="Tulare"
 /db_xref="taxon:51240"
 /clone="WSC0002_IVF_A11"
 /sex="Hermaphrodite"
 /dev_stage="Mid season fruit collected Aug 1"
 /lab_host="XLI0-Gold"
 /clone_lib="Seed coat
 collected Aug 1"
 /note="Organ: Seed coat; Vector: pTriplex2; Site 1: SfIIA; Site 2: SfIIB; Walnut nut samples were harvested from Tulare trees growing in the 'Stuke Block', in the Wolfskill experimental orchard located in Winters, California (USA).

Samples were harvested on August 1, 2001 between 8 and 10 am. Samples were then dissected the same day. Seed coat (pellicle) tissue was separated from embryos and frozen immediately in liquid nitrogen and stored at -80C. A gram of sample was removed and ground to a fine powder in liquid nitrogen. Total RNA was extracted using the hot borate procedure. Poly A+ RNA was obtained using the Poly(A) Purist kit (Ambion). The cDNA library was constructed using the SMART cDNA library kit (Clontech). Primary library was en masse evicted and plasmid DNA containing the cDNA library was isolated from the resultant bacterial population. Plasmid DNA was then transformed into ultra competent E coli cells (Xl10 Gold; Stratagene). Transformants were plated out on Q-trays (2000 cfu/tray), picked using a Qbot and archived in 384 well dishes."

ORIGIN

Query Match 9.7%; Score 279.6; DB 7; Length 766;
Best Local Similarity 60.3%; Pred. No. 2.8e-58;
Matches 462; Conservative 0; Mismatches 304; Indels 0; Gaps 0;

QY 1368 GGGGACAGTCTCTAGATGGCAGCAATGGACTACTATCACTCGGGCAATGGAAC 1427
DB 1 GGGGAAACGGGACCTGCTGATGGAAACGGTGGCAAACTTGTACTCCAGCAATACCAA 60

QY 1428 TTGCACCGCATGTTGGCGCAGCAGAGAAGATTTCTATGAATGTGGCTTTTGGGCAAA 1487
DB 61 ATGAAGTTAATGCTGCAATCACTGATGAGATGCCCTTAATTTTGGTTTACGGGAAA 120

QY 1488 GGCAATAGCTCTAGCAAAAACAACTTGTAGAAACAAGTAGAAGCGGCGCAATTTGTTTT 1547
DB 121 GGGACAGTGCAAAACCTGATGACTACATGAATAATTAGACGCGGCAATGGGACTG 180

QY 1548 AAATTCATGAAGACTGGGGGCAACACCAAGTGGCATGATCACTGTGTTAGCGTGGCA 1607
DB 181 AAGCTGCATGAGCACTGGGGAACCTACTCTGTGCAATAGACAAATGCTTTGAATGTGCA 240

QY 1608 GATGAATAGATGCAAGTTTGTATCCACCGCATACAGTCAATCAGCAGCGGTATGTA 1667
DB 241 GAACAATATGATCCAGGTTAATATCCACAGGACACCTTGAATGAATCGGGATTTGTT 300

QY 1668 GATGACACCTTAAATGAACGGGCGCGCATCCATCGCTTACCACATTTGAGGAGCG 1727
DB 301 GAACATTTCTATGCTGCAATTAAGGAAGAATAATTCATACCTACCACAGTGAAGTGT 360

QY 1728 GGTGGAGGACACTCACTGATGTATFACCAATGGCAGGCGAGCTCAATATTTACCCCTCC 1787
DB 361 GGTGGTGTTCATGCTCCAGATATCAATTAAGATATGTGGTGTAAACATGTCTGCCATCA 420

QY 1788 TCCACACCCCACTATTCCTTATACCATTAATACGGTTGCAAGAACACTTAGACATGCTC 1847
DB 421 TCTACGAACCCCAACAGCTTTTACTTTCAACACTATATAGATGAGCATCTTGACATGCTG 480

QY 1848 ATGACATGCCACCTAGACAAAACATCCGGAGGATTTACAAATTTTCTCAAAGCCGT 1907
DB 481 ATGGTCTGCATCACTCGATTAAGATATTCGAAGAGAGCTTGCTTTTGTGATCAAGG 540

QY 1908 ATCCGCGCCGGCTCTATCCGGCTGAAGATGTGCTCCATGATATGGGTGTGATCGCGATG 1967
DB 541 ATAAGGGCTGAAACAATTCCTGCAGAGGATATTTTGCATGATATGGGGCAATTAGCATC 600

QY 1968 ACAGCTCGATTCGCAAGCAATGGGCGTGCAGGGCAAGTGTCTTCTGAACTTGGCAG 2027
DB 601 ATTGCTTCGATGCAACAGCTATGGGTGCTGCTTTGGAGAGGTGATTAGCAGAACTTGGCAA 660

QY 2028 ACTGGCGATAAGAAATAAAGAAATTTGGTAAGCTTCTTGAAGATGCGCAAGATACGAT 2087
DB 661 ACTGCCCAACAGATGAAGTTACAAGAGGGGTGCAATTTGACCTTAGTGGACACAGCATGAC 720

QY 2088 AATTTCCGCAATTAAGCGCTACATCTCCAATATACACTATCAACCCCG 2133
DB 721 AATCTTCGTATCAAGGTTTACATGCTAATATACAAATAAATCCTG 766

RESULT 11
CF826586/c

LOCUS
DEFINITION

CF826586 986 bp mRNA linear EST 01-APR-2004
EST703988 Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb Coccidioides posadasii cDNA clone CIDB361 3' end, mRNA sequence.

ACCESSION
VERSION

CF826586
CF826586.1 GI:45932643

KEYWORDS
SOURCE

Coccidioides posadasii
Coccidioides posadasii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Onygenales; mitosporic Onygenales; Coccidioides.
1 (bases 1 to 986)
Gardner, M.J. and Cole, G.T.
Analysis of gene expression in Coccidioides posadasii mycelia and spores via expressed sequence tags
Unpublished (2003)
Other ESTs: EST703969
Contact: Gardner MJ
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org.

REFERENCE
AUTHORS

TITLE

JOURNAL
COMMENT

FEATURES
source

1. 986
/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/clone="CIDB361"
/dev_stage="saprobic phase (mycelia)"
/lab_host="E. coli DH10B, T1 phage resistant"
/clone_lib="Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb"
/notes="Vector: pExpress 1; Site 1: Not 1; Site 2: Eco RV; Coccidioides posadasii saprobic phase cDNA library, size fractionated cDNA 2 to 4 kb"

ORIGIN

Query Match 9.5%; Score 274.2; DB 6; Length 986;
Best Local Similarity 59.1%; Pred. No. 6.6e-57;
Matches 512; Conservative 0; Mismatches 343; Indels 12; Gaps 2;

QY 1721 GGGAGCGGTGGAGGACCTCACTGATGTTATCACCATGGCAGCGAGTCAATATCT 1780
DB 986 GGGTGTGGAGGAGGCCACGCTCCAGATATCATCGTCTCGAGAAGCCAAACGCTCT 927

QY 1781 ACCTCTCTCACACCCCACTATTCCTATACCATTAATACGGTTCAGACACTTAGA 1840
DB 926 GCCCAGCAGTACGAATCCCACTCGTCCGTATACGGTAAATACTTTAGATGAACATCTGGA 867

QY 1841 CATGCTCATGACATGCCACCTAGACAAACCATCCGCGAGATTTTACAATTTCTCA 1900
DB 866 CATGGTAATGCTGCGCATCATTTGTCCAAAGATATTCCTGAAGACGTTGGCTTTCGGA 807

QY 1901 AAGCCCGTATCCGCGCGCTCTATCGCGGTGAAGATGTCTCCATGATATGGGTGTGAT 1960
DB 806 AAGCCGATCCGATCCGAGACAATTCCTCAGAAGACGTTCTTCATGACGCGGAGCAT 747

QY 1961 CGGGATGACAAGCTCGGATTCGCAAGCAATGGGGGTGAGGCGAAGTATTCCTCGAAC 2020
DB 746 CAGCATGTATCTCCGCACTCTCAAGCTATGGGACGCTGTGGAGAGTTGTGTTCGGAC 687

QY 2021 TTGGCAGCTGGGATAAGATAAAGAAATTTGGTAAGCTTCTCTGAAGAT---GCCAA 2077
DB 686 ATGGAACTGCACTGCAATAAGATAAATGGAAACGAGGCGGACTCAAGGAAGATGAAGGAC 627

QY 2078 AGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTATCAACCCCGTTT 2137

Db 626 GGATTCGTGATAATTTTAGGGTTAAACGGGTATATATCAGCAAGTACACCATCAACCCCTGCCAT 567
QY 2138 GACCCACGGCGTGAGAGGATATATCGGCTCTGTGGAAGGGCAAGATCGCCGACTTGGT 2197
Db 566 TGCACAGGGGATGGCCACACTATTGGGAGCGTGGAAAGTTGGCAAGACCGCTGATTGGT 507
QY 2198 GGTGTGGAATCCGCTTTTGGCGGTAAACCCAAATCGTGATCAAAAGGCGGTATGGT 2257
Db 506 TCTGTGGAATTTGCCAACTTTGGGACTAAACCGAGTATGGTCTTGAAGTCTTGAATGGC 447
QY 2258 GGTCTCTCTGAATCGGCGATTTAAACGGCTCTGTGCCACTCCCCACCGGTTATTA 2317
Db 446 TGTCTCAGCGCAGATGGGTATCCCAATGGCTCTATCCCAATCGAGCTTATATAT 387
QY 2318 CCGGAAATGTTTGGGCATCACGGCAAGGCGAAATTTGACACAGCATCACTTTTGTTC 2377
Db 386 GAGGCTATGT-----ACGTAGTCTTAAACCTTAAAGCTCAATCATGTTTCGTATC 336
QY 2378 CAAAGTCGCTATGAATAAGCGGTGAAGAAAGCTGGGCTTAGAGCGCAAGTTCTACC 2437
Db 335 CCAAGCATCCATCAAGCTTGGTATCATCGACAGTTACCACTGAAGAAGCGGATCGAGCC 276
QY 2438 GGTCAAAATCGCGTAAACATCAACGAAGAGCTTCAAGTTCAACGACAAACGGCAA 2497
Db 275 AGTGAAGAATTTGCGAATAATAAGCAAGAGATATGAATTTAATGATATTATGCCAA 216
QY 2498 AATCACCCTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAACTCTGCACCTTAA 2557
Db 215 AATGAGAGTCGATCCGAGAGCTATGTTGTCGAGGCTGACGGGAGAGTGCACCGCTGA 156
QY 2558 ACCACCTCGAAGTGCCTTAGCCCCA 2584
Db 155 GCGAGTCAGAGTTGCTTTAAACACA 129

RESULT 12
AZ935182/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
COMMENT
FEATURES
source
ORIGIN

AZ935182 784 bp DNA linear GSS 24-APR-2001
BJ Ba0003020f B. japonicum BAC library Bradyrhizobium japonicum
genomic, genomic survey sequence.
AZ935182
GSS.
Bradyrhizobium japonicum
Bradyrhizobium japonicum
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Bradyrhizobiaceae; Bradyrhizobium.
Tomkins,J.P., Wood,T.C., Stacey,M.G., Loh,J.T., Judd,A.,
Goicoechea,J.L., Stacey,G., Sadowsky,M.J. and Wing,R.A.
A marker-dense, sequence-ready map of the Bradyrhizobium japonicum
genome
Genome Res. 11 (8), 1434-1440 (2001)
11483585
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Class: BAC ends
High quality sequence stop: 719.
Location/Qualifiers
1. 784
/organism="Bradyrhizobium japonicum"
/mol_type="genomic DNA"
/strain="USD110"
/db_xref="taxon:375"
/lab_host="E. coli"
/clone_lib="B. japonicum BAC library"
/note="Vector: pindigo536; Site_1: HindIII"

Query Match 9.5%; Score 273; DB 9; Length 784;
Best Local Similarity 64.0%; Pred. No. 1.2e-56;
Matches 458; Conservative 0; Mismatches 253; Indels 5; Gaps 3;
QY 1468 ATCTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAACAACTTGTAGAACAGTAG 1527
Db 715 ATCTGGCATTTTCGGGCAAGGCAACGCCCTCGCGCCCGCGCTGGTCAGATGATCA 656
QY 1528 AAGCGGCGCATTTGGTTTAAATTTGCATGAAGACTGGGCGACAAACCAAGTGCATCG 1587
Db 655 AGNCGGCGCATCGCGCTGAAGCTGCAAGGATTTGGGCGAC-ACGCGGCGCGATCG 597
QY 1588 ATCACTGCTTTAGCGTGGCAGATGAATAAGATGCAAGTTTGTATCCACACCGATACAG 1647
Db 596 ACAATGCTGTGGTGGCGGACGATTTACGACATCCAGGTCAATGATCCACACCGATAGC 537
QY 1648 TCAATGAGGAGGTATGTAGATGACACCTTAATGCAATGAACGGGCGGCATCCATG 1707
Db 536 TGAAGATCCGGGCTTCGTGAGGATACGATCAAGGGCTTCAAGGGCGGCAACCATCCAG 477
QY 1708 CCTACCAATTTAGGGAGCGGCTGGAGGACACTCACCTGATGTTTATCACCATGGCAGCG 1767
Db 476 CTTTCCACACCGAGGGCGCGCGCGGTCAAGCCCGGACATCATCAAGGTTCGAGGGC 417
QY 1768 AGCTCAATATTTCAACCTCTCCACACCCCACTATTTCCTATACCAATTAATAGGTTG 1827
Db 416 TGAAGAACGTGCTGCCGTCTCGACCAACCGAGCGGCCCTTCAACCGCAACCATCG 357
QY 1828 CAGAACATTTAGACATGCTCATGACATGCCACCACTAGACAAAGGATCCGCGAGGATT 1887
Db 356 ACAGCATCTCGACATGCTGATGTTGTGCCACCACTCGATCCCTCGATCGCGAAGATC 297
QY 1888 TACAATTTTCTCAAAGCGGTATCGCCCGGCTCTATCGCGCTGAAGATGTGCTCCATG 1947
Db 296 TTGGTTTCGCGAAGCCGTATCCGCAAGGAGACCATCGCGCGGAGACATCCTGCAGC 237
QY 1948 ATATGGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGG-GCGTGCAGGCGAA 2006
Db 236 ATCTGCGCGCTCTCGATGATGCTCTCGNACTCCAGGCCATGGGCCGCTCGCGAA 177
QY 2007 GTGATTCCTCGAATTGGCAGACTGGCGATGAAGATTAATAAAGATTTGGTAAGTTCC- 2065
Db 176 GTCATCATCGGACCTGGCAGACCGCGACAAAGATGAAGACGCGCGGATCGCTGCCG 117
QY 2066 --TGAAGATGCAAGATAAGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACT 2123
Db 116 CAGGCAAGGCAAGGACAAAGCAATTTCCGCGTCAAGCGCTACATCGCCAAATACAG 57
QY 2124 ATCAACCCCGCTTTGACCCACGCGGTGAGCGAGTATATCGGCTCTGTGGAAAGAGG 2179
Db 56 ATCAACNCGGATCGCGCACGGCGTGTGGAAGCTGATCGGTTCCGTCGGAAGAGG 1

RESULT 13
CO027797/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CO027797 990 bp mRNA linear EST 10-JUN-2004
EST806181 Coccidioides posadasii spherule cDNA library, 0.5 to 5.3
kb Coccidioides posadasii cDNA clone CIPAM57 3' end, mRNA sequence.
CO027797
EST.
Coccidioides posadasii
Coccidioides posadasii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygena; mitosporic Onygenales; Coccidioides.
Gardner,M.J. and Cole,G.T.
Analysis of gene expression in Coccidioides posadasii mycelia and
spherules via expressed sequence tags
Unpublished (2003)
Other ESTs: EST806182
Contact: Gardner MJ

The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@igr.org

FEATURES

Location/Qualifiers
i. 990

/organism="Coccidioides posadasii"

/mol_type="mRNA"

/strain="C735"

/db_xref="taxon:199306"

/clone="CIPAM57"

/dev_stage="spherules"

/lab_host="E. coli DH10B, T1 phage resistant"

/clone_lib="Coccidioides posadasii spherule cDNA library,
0.5 to 5.3 kb"

/notes="Vector: pExpress 1; Site_1: Not I; Site_2: Eco RV;
Coccidioides posadasii spherule cDNA library, 0.5 to 5.3
kb"

ORIGIN

Query Match 9.1%; Score 262.6; DB 7; Length 990;
Best Local Similarity 58.8%; Pred. No. 5.2e-54;
Matches 525; Conservative 0; Mismatches 339; Indels 29; Gaps 3;

1696 GCGCCATCCATGCTACCAATGAGGAGCGGTGGAGGACACTCACCTG-ATGTTATC 1754
Db 990 GAGTGATTACATGCTACCACTGAGGTGCTGAGGAGGCCAGCTCCAGATATCATA 931

1755 ACCATGGAGCGAGTCAATATTCCTCTCCACCAACCCCACTATTCCTATACC 1814
Db 930 TCCGTGCTGAGAAGCAACGCTCTGCCAGCAGTACGAATCCCACTCGCTGATACG 871

1815 ATTAACTCGTTGCAGAACACTTAGCATGCTCATGACATGCCACCACTAGACAAACGC 1874
Db 870 GTAATACTTTAGATGAACATCTGGACATGTTAATGCTCTGCCATCATTTGTCCAAAGAT 811

1875 ATCCGCGAGGATTTACAATTTCTCAAAGCGGTATCGCCCGCGCTCTATCGCGCTGAA 1934
Db 810 ATTCTGAGACGTGCTTTTGGGAAGCGCGATCCGATCCGAGACATTTGCTGCAGAA 751

1935 GATGCTCTCATGATATGGTGTGATGCGGATGACAAAGCTCGGATTCGGAAGCAATGGGG 1994
Db 750 GACGTTCTTCATGACACGGAGCATCAGCATGCTATCTCTCGGACTCTCAAGCTATGGGA 691

1995 CGTGAGGCGAGTATTCCTGCACTTGGCAGACTGCGGATGAAGATATAAAGAAATTT 2054
Db 690 CGCTGTGGAGAAGTTGTTTCGGACATGGAACACTGACATAGAATAAATGGAACGA 631

2055 GGTAAAGCTTCTCAAGAT--GGCAAGATAAGATAATTTCCGATTTAAGCGCTACATC 2111
Db 630 GGGCGACTCAAGGAATGAGGAGCGGATTTGATAAATTTTAGGGTTAAACGGTATATC 571

2112 TCCAAATACATCATCAACCCCGCTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTG 2171
Db 570 AGCAAGTACACCATCAACCTGCGCATTCGACAGGGGATGCCACACTATTTGGAGCGTG 511

2172 GAAGAGGCAAGATCCCGACTTGGTGTGTGGAATCTCTGCTTTTGGCGTAAACCC 2231
Db 510 GAAGTTGGCAAGACCGCTGATTTGGTTCTGTGGAATTTGGCAACTTTGGGACTAAACCG 451

2232 AAAATCGTATCAAGCGGTATGGTGTCTCTGTAATGGCGGATTTCTAACGGCTCT 2291
Db 450 AGTATGGTCTTGAAGTCTGGAATGGCTGTCTCAGCGCAGATGGGTGATCCCAATGGCTCT 391

2292 GTGCCCACTCCCAACCGGTTTATTACCGGAAATGTTTGGGCGATCACGGCAAGGGAAA 2351
Db 390 ATCCCAACATCGAGCTATTTATTAGAGCCTATGTAGCTCTCTCA-----344

2352 TTTGACACCGACATCACTTTGTTTCCAAAGTGGCTATGAAATGGCGTGAAGAAAG 2411
Db 343 -----ATCATGTTCTGTTATCCCAAGCATCCATCAAGCTTGGTATCATCGACGT 296

Qy 2412 CTGGGCTTAGAGCCCAAGTTCTACCGTCAAAAACCTGCCGTAAACATCACCAAGAAAGAC 2471
Db 295 TACCACCTGAAGAAGCGGATCGAGCCAGTGAAGAAATTTGTCGGAATATAAGCAAGAGAGAT 236

Qy 2472 TTCAAGTTTCAACGACAAAACGGCAAAATACCGTCCGATCCGAAACCTTCGAGGTCTTT 2531
Db 235 ATGAAATTTAATGATATTATGCCCAAAATGAGATCGATCCGAGAGGTATTTGTCGAG 176

Qy 2532 GTAGATGGCAAACTCTGCACCTCTTAAACCCCACTCGCAAGTGCCTCTAGCCCA 2584
Db 175 GCTGACGGGAGAGTGCACCGCTGAGCCAGTGTACAGATTCCTTTAAACACA 123

RESULT 14
CN907928
LOCUS
DEFINITION
CN907928
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

684 bp mRNA linear EST 07-JUN-2004
0310109ABLC001286HT (ABLC) Braeburn cell culture three days after
subculture Malus x domestica cDNA clone ABLC001286, mRNA sequence.

Malus x domestica
Malus x domestica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.

1 (bases 1 to 684)
Beuning, L., Bowen, J., Crowhurst, R., Gleave, A., Janssen, B.,
McArtney, S., Newcomb, R., Ross, G., Snowden, K., Walton, E. and Yauk, Y.
HortResearch Apple EST Project
Unpublished (2004)
Contact: Gleave, A.
Sequencing Facility
The Horticulture and Food Research Institute of New Zealand Ltd
120 Mt Albert Rd, Mt Albert, Auckland, New Zealand
Tel: 00 64 09 815 4200
Fax: 00 64 09 815 4201
Email: est@hortresearch.co.nz.

Location/Qualifiers
1. 684
/organism="Malus x domestica"
/mol_type="mRNA"
/db_xref="taxon:3750"
/clone="ABLC001286"
/tissue_type="Cell cultures"
/dev_stage="Exponential phase of growth (three days after
subculture)"
/clone_lib=" (ABLC) Braeburn cell culture three days after
subculture"
/notes="Vector: pBluescript SK(-); Library sequenced by
Genesis Research & Development"

ORIGIN

Query Match 9.1%; Score 261.2; DB 7; Length 684;
Best Local Similarity 61.4%; Pred. No. 1.1e-53;
Matches 419; Conservative 0; Mismatches 263; Indels 0; Gaps 0;

Qy 1546 TTTAAATTGCATGAAGACTGGGGCACAACACCAAGTGGGATCGATCCTGTTGAGCGTGG 1605
Db 3 TGAAGCTGCATGAGGACTGGGGAAAGCACTCTCTGTCGAATCGACAATTTGTTTGGCTGTG 62

Qy 1606 CAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGCGAGTTATG 1665
Db 63 GAGATCAATATGACATCCAGGTTAACTATCCAGACACTTGAATGAATCTGGATTG 122

Qy 1666 TAGATGACACCTTAAATGCAATGAACGGCGCGCCATCCATGCTTACCACATTTAGGGAG 1725
Db 123 TAGAGCATACAATGTGTCATTTAAGGAAGAACTATCCATACTTACCACAGTGAAGTG 182

Qy 1726 CGGCTGGAGGACACTCACTGATGTTATACCAATGGCAGCGGAGCTCAATATTTACCT 1785
Db 183 CAGGTGGGGGCCATGCTCCAGATATAATCAAGTCTGCGGTGTGAAAAAATGTCCTGCCAT 242

```

QY 1786 CCTCCACCCACCCACTATTCCTATACCAATTAATACGGTTGCGAACAACCTTAGACATGC 1845
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
243 CATCAACGAATCCCAACCGCCCTTTCACCTCCAAATACATATAGATGAGCATCTTGTATATGC 302
QY 1846 TCATGACATGCGCACCCACCTAGACAAACGCAATCGCGAGGATTTACAAATTTCTCAAAGCC 1905
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
303 TGATGGTTGGCATCACCTTGACAGGACATTCAGNAGATGTAGCTTTTGTCTGAATCAA 362
QY 1906 GTATCCGCGCGGCTCTATCGCGCTGAAGATGTCTCCATGATATGGGTGTGATCGCGA 1965
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
363 GAATAAGGGCTGAACAATTTGTCAGAGATATTTTGCACGATATGGGGGCCATTAGCA 422
QY 1966 TGACAGCTCGGATTCGACAGCAATGGGGCTGACGCGAGTGTATCTCGAACTTGGC 2025
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
423 TTGTATCTTCTGATTCACAGGCTATGGGTGCATTTGGAGAGGTGATAATCAGAACTTGGC 482
QY 2026 AGACTCGGATGAAGATTAATAAGAAATTTGGTAAAGCTTCTGAAGATGCGCAAGATAACG 2085
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
483 AAGAGCTGACAGATGAATACAAAGAGGTGATAGAACCTAGTGGATCCGACATG 542
QY 2086 ATAATTTCCGATTAAGCGCTACATCTCCAATACACTATCAACCCCGCTTTTGACCCACG 2145
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
543 ACAATTTCCGGATCAACAGATACATCGCAAAGTACACTATAAATCCAGCAATAGTTAATG 602
QY 2146 GCGTAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCGACTTGGTGTGTGGA 2205
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
603 GATTTTCTCATGATGTGGTCTTATTGAGTGGGAGCTGGCTGATCTTTGTTCTATGGA 662
QY 2206 ATCCTGCCTTTTGGCGTAAA 2227
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
663 AGCAATGTTCTTTGGCGCAA 684

```

RESULT 15

CF714848/c

LOCUS CF714848 921 bp mRNA linear EST 16-AUG-2004
 DEFINITION C.neoformans strain JEC21 Cryptococcus neoformans var.
 neoformans cDNA clone CCAF343, mRNA sequence.

ACCESSION

CF714848

VERSION

CF714848.1

KEYWORDS

EST.

SOURCE

Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
 var. neoformans)
 Cryptococcus neoformans var. neoformans
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
 Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
 Filobasidiella.

REFERENCE

1 (bases 1 to 921)

Loftus, B.

End sequencing of clones from a Full length enriched, normalized

JEC21 cDNA library

Unpublished (2003)

Other_ESTs: CCAF343TR

Contact: Brendan Loftus

The Institute for Genomic Research (TIGR; www.tigr.org)

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-3543

Fax: 301-838-0208

Email: bjloftus@tigr.org

Seq primer: TF.

Location/Qualifiers

1..921

/organism="Cryptococcus neoformans var. neoformans"

/mol_type="mRNA"

/db_xref="taxon:40410"

/clone="CCAF343"

/note="Vector: pCMVSPORT6; Site 1: NotI EcoRV; The full

length, normalized library was prepared from a variety of

conditions using RNA provided by Joseph Heitman and

Jennifer Lodge"

Query Match 9.1%; Score 261; DB 6; Length 921;
 Best Local Similarity 55.4%; Pred. No. 1.3e-53;
 Matches 504; Conservative 0; Mismatches 405; Indels 0; Gaps 0;

```

QY 1694 GCGCGCATCCATCGCTTACCACATTTAGGGAGCGGGTGGAGGACACTCACCTGATGTTAT 1753
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
921 GAGACATTTTACAGTTTACCAGTCCAGACTGAGAGAGCTGTGTGGGAGATGTCAGACATCAT 862
QY 1754 CACCATGGCAGGGAGCTCAATATTTCTACCCCTCTCTCCACCACCCCTCACTATTCCTATPAC 1813
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
861 CGTTGTCTGTGAATATGAAAAACGTTTCCCTAGTTCCACCAACCACTAGGCTTATGC 802
QY 1814 CATTATACGGTTTGCAGAACACTTACGATCTCATGACATGCCACCACTAGACAAACG 1873
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
801 TGTAAACACTCTTTGATGAGCATCTTGACATGCTTATGTTCTGTCCACCACTCGATAAGTC 742
QY 1874 CATCCGCGAGGATTTACAAATTTTCTCAAACCGGTATCCGCCCGGCTCTATCGGGCTGA 1933
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
741 TATCCCGGAGGACATTCGCTTTGCCGACTCTCGTATCCGTTCTGAAACCGTTGCAGCCGA 682
QY 1934 AGATGTCTCCATGATATGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGG 1993
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
681 AGAGTTTTTACAGGACACAGCGCGATTTCCATGATCTCATCTGACTGTCAAGCTATGGG 622
QY 1994 GCGTGACGGCAAGTGAATCTCTGAACTTGGGAGACTGCGGATAGAAATAAAAAAGAAAT 2053
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
621 TCGTATCGGTGAATCATCACTCGTATCATGCGGTACGGCTGCGAAGATGAAGCAATTCGG 562
QY 2054 TGGTAGCTTCTTGAGATGCGAAAGATTAACGTAATTTCCGCAATTAAGCGCTACATCTC 2113
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
561 TGGTCTCTTGAAGGCGATGAACCCACGAGAGACAATAGGGGTGAACCGTATGTTGC 502
QY 2114 CAAATACACTATCAACCCCGCTTTGACCCACGGGTGAGCGAGTATATCGGCTCTGTGGA 2173
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
501 CAAGTACACTATTAACCCCGCATTACCACGGTATGTGCGACCTCATCGGTCAAGTTGC 442
QY 2174 AGAGGGCAAGATCGCGACTTGGTGTGGAATCTCTGCTTTTGGCGGTAAACCCAA 2233
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
441 CGTGGGCTGTCTGGCAGATCTGGTCTTCTGGAACGGCAGAGTCTTTGGTGTAGACCCGA 382
QY 2234 AATCGTATCAAAGGCGGTATGGTGTCTTCTGAAATGGGCGATTTCTAAGCGTCTGT 2293
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
381 GATGATCTCAAGGGTGTGTCTATGTTGGCTGCAATGGGCGATGCGAACGCTTCGAT 322
QY 2294 GCCCACTCCCAACCGGTATTTATACCGGAAATGTTTGGGCATCAAGCAAGGCAAAAT 2353
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
321 CCCCACTGTGAGCCCGTCAATGGCAGCCCATGTGGGGCTCTCAGCCTGAGGCGCTGC 262
QY 2354 TGACACCAAGCATCACTTTTGTTCCTCAAGTGCCTATGAAATGGCGTGAAGAAAGCT 2413
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
261 ACTCAATTCATTTTGGGTGAGCCAGGCTCTCTTGACAAGGATCTCGTGAAGAGATT 202
QY 2414 GGGCTTAGAGCGCAAGTTCTACCGGTCAAAAACCTGCGGTAAACATCAACAGAAAGACTT 2473
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
201 CAACATCAAGAAGAGGGCGGAGGCTGTCAAGAACTGTGCTCAATTTGAAAGAGGATAT 142
QY 2474 CAAGTTCAAGCAAAAACGCGCAAAAATCACCGTCGATCCGAAACCTTCGAGGTCTTTGT 2533
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
141 GAAGTGAATGACAGATGCGGAAGATGCTGTGATCCGGAGACTTACAGCTTACAGCTTACGC 82
QY 2534 AGATGGCAAACTCTGCACCTCTAAACCCACCTCCGCAAGTGGCTCTTAGCCCGCGCTACAC 2593
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
81 TGACGGCGCTCTTTGCGACGCTCCACCGGACAGCAAACTCCCACTGACCAAGAGATCTT 22
QY 2594 TTTCTTCTA 2602
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
21 CGTTTACTA 13

```

Search completed: November 29, 2005, 07:20:55

Job time : 10217 secs

ORIGIN

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2005, 00:04:09 ; Search time 522 Seconds
(without alignments)
9817.453 Million cell updates/sec

Title: US-09-904-994B-1

Perfect score: 2883

Sequence: 1 rrgagattttccarctt.....aaaaagtagagccacagg 2883

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq.*
- 7: /cgn2_6/ptodata/1/ina/pp COMB.seq.*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
- 9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1075.2	37.3	2735	2	US-08-920-095-1
2	1075.2	37.3	2735	6	PCT-US96-05800-1
3	1062.6	36.9	2619	2	US-08-467-822-19
4	1062.6	36.9	2619	3	US-08-432-697-19
5	1062.6	36.9	2619	3	US-08-466-248-19
6	1058.6	36.7	4824	3	US-09-431-705-1
7	1058.6	36.7	4824	3	US-09-431-705-19
8	737.6	25.6	6131	2	US-07-732-242C-8
9	693.4	24.1	2400	2	US-08-967-513-1
10	693.4	24.1	2400	2	US-08-687-645B-1
C 11	664.4	23.0	1830121	3	US-09-557-884-1
C 12	664.4	23.0	1830121	3	US-09-643-990A-1
C 13	664.4	23.0	1830121	3	US-10-158-865-1
C 14	662.8	23.0	5966	3	US-08-956-171B-22
C 15	662.8	23.0	5966	3	US-08-781-986A-22
16	653	22.7	1710	3	US-09-543-681A-1857
17	630.6	21.9	8729	3	US-09-453-702B-258
18	630.6	21.9	8729	3	US-10-114-170-258
19	630.6	21.9	87563	3	US-09-453-702B-57
20	630.6	21.9	87563	3	US-10-114-170-57
21	623.4	21.6	1716	6	US-08-487-429A-3
22	623.4	21.6	1716	6	PCT-US96-05320A-541
23	617.8	21.4	1878	3	US-09-489-039A-2045
24	612	21.2	3164	3	US-09-710-279-3828

25	612	21.2	3234	3	US-09-710-279-3853	Sequence 3853, Ap
26	578.6	20.1	1716	3	US-09-710-279-117	Sequence 117, App
27	578.6	20.1	1722	3	US-09-134-001C-2189	Sequence 2189, Ap
28	578.4	20.1	1704	3	US-09-328-352-1786	Sequence 1786, Ap
29	572.6	19.9	1701	3	US-09-252-991A-10316	Sequence 10316, A
C 30	557	19.3	1653	3	US-09-252-991A-10546	Sequence 10546, A
31	551.8	19.1	1833	3	US-09-602-777A-13	Sequence 13, Appl
32	528.4	18.3	1625	3	US-09-602-777A-15	Sequence 15, Appl
C 33	512	17.8	1797	3	US-09-601-198-89	Sequence 89, Appl
34	502	17.4	4403765	3	US-09-103-840A-2	Sequence 2, Appl1
35	502	17.4	4411529	3	US-09-103-840A-1	Sequence 1, Appl1
36	424.4	14.7	3475	3	US-09-710-279-4032	Sequence 4032, Ap
37	356.6	12.4	951	3	US-09-252-991A-10513	Sequence 10513, A
C 38	213.2	7.4	636	3	US-09-252-991A-10653	Sequence 10653, A
39	210.2	7.3	684	3	US-10-012-819-153	Sequence 153, App
40	173.2	6.0	363	3	US-10-012-819-159	Sequence 159, App
41	101.6	3.5	613	3	US-09-533-559-7520	Sequence 7520, Ap
42	100.6	3.5	462	3	US-09-543-681A-1913	Sequence 1913, Ap
43	99.2	3.4	312	3	US-09-328-352-1761	Sequence 1761, Ap
44	97.4	3.4	300	6	PCT-US96-05320A-543	Sequence 543, App
C 45	97.4	3.4	309	3	US-09-710-279-121	Sequence 121, App

ALIGNMENTS

RESULT 1
US-08-920-095-1
; Sequence 1, Application US/08920095
; Patent No. 5837240
; GENERAL INFORMATION:
; APPLICANT: Cynthia K. Lee et al.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,095
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US/08/431,041
; APPLICATION NUMBER: 28-APR-1995
; FILING DATE: 28-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2735 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-920-095-1

Query Match 37.3%; Score 1075.2; DB 2; Length 2735;
Best Local Similarity 66.0%; Pred. No. 5.8e-267;
Matches 1622; Conservative 0; Mismatches 798; Indels 38; Gaps 3;

QY 182 AAAATTTAAACAGAGTAATAGGTGAAACTCACACCCAAAGAGCAAGAAAGTTCTTG 241
Db 78 ACACITTTAAGAATAGAGATGAGATGAACTCACCCAAAGAGTTAGATAAGTTGATG 137
QY 242 TTATATATTATGCGGCGAAGTGGCTAGAAAGCGCAAGACAGAGGCTTAAAGCTCAACAA 301
Db 138 CTCACCTACGCTGAGAAATGGCTTAAAAACGCAAGAAAGACATTAAGCTTAACAT 197
QY 302 CCGAAGACCATTTCTTACATTAAGTGGCCATATATTATGACGAAGCGCCGCTTGAAGAAAAA 361
Db 198 GTAGAAGCAGTAGCTTTGATTAGTGCCCATATATTATGGAAGAGAGAGCTGTTAAAG 257
QY 362 ACCGTTGCCAGCTTATGGAAGAGTGCATGCACATTTTGTGAAGAAAGATGAAGTAATGCC 421
Db 258 ACTGCGGCTGAATTTGATGCAAGAGGCGCACCTCTTTTAAACCCAGATGATGTGATG 317
QY 422 GGGGTGGGTAATATGTTCCGATCTAGGTGTAGAGCAGCCATTTCTGATGTTAGCAAA 481
Db 318 GGGGTGGCAAGCATGATCCATGAGTGGGTATTTGAAGCGATGTTTCTGATGGACATAA 377
QY 482 CTTGTAACTGTGAATTTGGCCCATCGAACCAAGATGAGCACCTTCAAAGCGGCGAAGTGAAA 541
Db 378 CTCGTAACCGTGCATACCCCTATTGAGGCCAATGGTAATTTAGTTCTGGTGAGTTG- 434
QY 542 TTTGGTTGCGATAAAGACATCGAGCTCAATGCGAGGCAAGAAAGTAAACCGAACTTGA 601
Db 435 TTCTTAAAAAATGAAGACATCACTATCAACGAAGGCAAAAAAGCCGTTAGCGTGAAGTT 494
QY 602 ACTAATGAGGGCTAAATCTTGGCATGTGGGTAGCCATTTCCATCTTTTGAAGCTAAC 661
Db 495 AAAATGTTGGGACAGACCGGTTCAAATCGGCTCACACTTCCATTTCTTTGAAGTGAAT 554
QY 662 AAGGCACTAAATTTGCAATCGTGAAGAACCCATGCGCAAGCCCTAGATATTTCCCTCTGG 721
Db 555 AGATGCTTAGACTTTGACAGAGAAACCTTTCGGTAAACGCTTAGACATTTGCGAGCGG 614
QY 722 AACACGCTAGCAGTTGGGCGAGACAAACCCGCAAGTGCAGTTGATTCCTCTTTGGTGG 781
Db 615 ACAGCGTGAAGATTTGAGCTGCGCAAGAAAAATCCGTAGAAATTTGATGACATTTGGCG 674
QY 782 AGTAAAAAGTGTGCGATGAAGCGGCTTTGTGATTAACATCGCGATGAAGCCATAAA 841
Db 675 AACAGAGAAATCTTTGGATTAAACGCAATTTGTTGATAGAACAGCAGACAAACGAAG 734
QY 842 CATTAAGCGCTTGCAAGGCGAAATCTCACGGATTT----- 877
Db 735 AAAATTCCTTACACAGAGCTAAAGACGCTGGTTTTTCATGGCGCTAAAGCGATGACAC 794
QY 878 -----ATCAAGTAAGGAGACTCCCATGAAA-----ATGAAAAAACAGAAATATGTAA 926
Db 795 TATGTAATAAAACAAATTAAGGAGTAAGAAATGAATAAGATTTAGCAGAAAAAGATAATG 854
QY 927 ACCTACGGACCCAAAGCGGATAAGAGTGGCTTAGAGATACCGATCTTTGGGCGAGAA 986
Db 855 ATGTATGGTCTCTACAGCGGATAAAGTGAAGTTGGGCGATACAGACTTGTATCGCTGAA 914
QY 987 GTAGAACATGACTATACCATCTATGGGGAAGAACTTAAATTTTGGCGGGTAAACATATC 1046
Db 915 GTAGAACATGACTATACCATTTTATGGCGAAGAGCTTAATTCGGTGGCGGTAAACCCCTA 974
QY 1047 CGTGAGGTTATGGGTACAGACAAATAGCCCTGATGAAAAACCCCTAGATTTTATGTCATCA 1106
Db 975 AGAAGGCAATGACCAATCTAAACCCCTAGCAAGAGAGTTGGATTTTAAATTTATCACT 1034
QY 1107 AACCGCATGATATTCGATACACCGGATTTTAAAGCGGACATTTGGGATTTAAAGACGG 1166
Db 1035 AACGCTTTAAATCGTGGATTTACCGGATTTTATAAAGCGGATATTTGGTATTTAAAGATGC 1094
QY 1167 AAAATCCATGGCATTTGCAAGGAGGAAACAGACATGCAAGATGGGTAAGCCCTCAT 1226
Db 1095 AAAATCGTGATTTGTAAGGCGGTAAACAAACATGCAAGATGGCGTTTAAAGCAAT 1154
QY 1227 ATGTCGTGGGTGTGGGCAAGAGCCTAGCAGGGGAAGGTATGATTTATTACGCTGGG 1286

Db 1155 CTTAGCGTAGGTCTCTACTAGAGCCTTAGCGGTTGAAGTTTGTATCGTAACCGCTG 1214
QY 1287 GGAATCGATTACACACCCACTTCTTCTCCAAACAAATTTCCCTACCGCTCTAGCCAA 1346
Db 1215 GGTATTGACACACATCCACTTCTTACCCCAACAAATCCCTACAGCTTTTGCAGC 1274
QY 1347 GGGTTTAAACCAATGTTTGGAGGCGCACAGGCTCTGTAGATGCAAGATGCGACTACT 1406
Db 1275 GGTGTAACAAACCATGATTGGTGGTGAACCGGCTCTGTATGGCACTTAATGCGACTACT 1334
QY 1407 ATCACTCCGGGCAAAATGGAATTTGCCGCAATGTTGGCGCAGCAGAGAGATTTCTATG 1466
Db 1335 ATCACTCCAGGCAAGAAATTTAAATATGATGCTCAGAGCGGCTGAAGAAATTTCTATG 1394
QY 1467 AATGTGGGCTTTTGGGCAAGGCAATAGCTCTTAGCAAAAAACAACTTTGTAGAAACAAGTA 1526
Db 1395 AATTTAGGTTTCTTGGCTTAAAGGTAAAGCTTCTAAACGATGCGAGCTTAGCCGATCAA 1454
QY 1527 GAAGCGGCGGATGTTGTTTAAATTTGATGAAGACTGGGGCAAAACCAAGTGGCATC 1586
Db 1455 GAAGCGGCTGCGATTGGCTTTAAATTTCAAGAGACTGGGGCACCACTCTCTTCTGCAATC 1514
QY 1587 GATCACTGCTTAGCGCTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACA 1646
Db 1515 AATCATGCGTTAGATGTTGGGCAAAATACGATGTCAAGTCTGCTATCCACACAGACT 1574
QY 1647 GTCAATGAGGCGAGTTATGTAGATGACACCTTAATGCAATGAACGGGCGGCCATCCAT 1706
Db 1575 TTGATGAGAGCGGTTGTGTAGAGACACTATGGCTGCTATTGTGAGCGCATATGCAC 1634
QY 1707 GCCTAACACATTTAGGAGCGGTTGAGGACACTCACTGTATGTTATCACCATGGCAGGC 1766
Db 1635 ACTTTCACACTGAAGCGCTGGCGGCGACACGCTCTGTATTTATTTAAAGTAGCCGCT 1694
QY 1767 GAGCTCAATATTTACCTCTCCACACCCCACTATTCCCTATACCATTAATACGGTT 1826
Db 1695 GAACACAACTTCTTCCGCTTCCACTAAACCCCACTTCCCTTCCCGTGAATACAGAA 1754
QY 1827 GCAGAACACTTTAGACATGCTCATGACATGCCACCACTAGACAAACGATCCGCGAGAT 1886
Db 1755 GCAGAGCACATGGACATGCTTATGTTGTGACCACTCTGGATTAAGAGATTAAGAGAT 1814
QY 1887 TTACAAATTTTCTCAAAGCGTATCCGCCCGGCTCTATCGCGGCTGAAGATGTGTCCAT 1946
Db 1815 GTTCAGTTCTGATTTCAAGGATCCGCCCTCAAAACCATTTGGGCTGAAGACACTTTTGCAT 1874
QY 1947 GATATGGGTGTGATCGGATGACAGCTCGGATTTGCGAAGCAATGGGCGGTGCGAGCGAA 2006
Db 1875 GACATGGGGATTTTCTCAATCACAGTTCTGACTCTCAAGCGATGGGCGGTGTTGGGTGAA 1934
QY 2007 GTGATTTCTCGAACTTTGGCAGACTGCGGATAAGAAATAAAAAAGAAATTTGGTAAGCTTCT 2066
Db 1935 GTTATCACTAGACTTTGGCAACAGCTGACAAACCAAGAAAGAAATTTGGCGGCTTGAAA 1994
QY 2067 GAAGATGCCAAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATFACACTATC 2126
Db 1995 GAAGAAAAAGCGGATACGACAACTTCAGGATCAACCGCTACTTGTCTTAAATFACACTT 2054
QY 2127 AACCCCGTTTGAACCGAGGCTGAGGAGTATATCGGCTCTGTGGAAAGAGGGCAAGATC 2186
Db 2055 AACCCAGCGATCGCTCATGGGATTTAGCGAGTATGTAGGTTTCAGTAGAAGTGGGCAAGTG 2114
QY 2187 GCCGACTTGGTGTGGAATCTGCTTTTTCGCGTAAACCCCAATCTGTATCAAA 2246
Db 2115 GCTGACTTGGTATTTGGGAGTCCAGCATCTTTTGGCGTGAACCCCAATGATCATCAAA 2174
QY 2247 GGGGTATGGTGTCTTCTCTGAAATGGGCGATTTCTAACCGGCTCTGTGCCCACTCCCCAA 2306
Db 2175 GGGGATTTATTTGGTTAAGCCAAATGGGCGATGGACGCTTCTATCCCTACCCACAA 2234
QY 2307 CCGGTTTATTAACCGCAAAATGTTTGGGATCACCGCAAGGCGAAATTTGACACCGAGATC 2366

Db 2235 CCGGTTTATTACAGAGAAATGTTGGCTCATCATGTGTAAGCTTAAATACGATGCAAAATC 2294
QY 2367 ACTTTTGTTCCAAAGTCGCCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAGCGC 2426
Db 2295 ACTTTTGTCTCAGCGGCTTATGCAAGGATTAAGAGAGATTAGGACTTGAAGA 2354
QY 2427 CAAAGTTCTACCGGTCAAAACTGCGGTAAATCATCAAGAGAGACTTCAAGTTCAACGAC 2486
Db 2355 CAAAGTTGCGGTAAATAATTGCAGAAATATCACTTAAAGAGACATGCAATTTCAACGAC 2414
QY 2487 AAAACGGCAAAATCACCGTCGATCGAAACCTTCGAGGTCTTTAGATGCGCAACTC 2546
Db 2415 ACTACTGCTCAATGGAAGTCAATCTGAAACTTACCATGTGTTGCGTATGGCAAGAA 2474
QY 2547 TGCACCTCTAAACCCACCTCGCAAGTGCCTCTAGCCCGAGCGGTACACTTTCTCTAGG 2604
Db 2475 GTAACCTTCTAAACGAGCAATGAAGTGAAGTGGCGCAACTCTTAGCATTTCTTAGG 2532

RESULT 2

PCT-US96-05800-1
; Sequence 1, Application PC/TUS9605800
; GENERAL INFORMATION:
; APPLICANT: OraVax, Inc.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05800
; FILING DATE: 23-APR-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,041
; FILING DATE: 28-APR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/568,122
; FILING DATE: 06-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/020001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2735 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US96-05800-1

Query Match 37.3%; Score 1075.2; DB 6; Length 2735;
Best Local Similarity 66.0%; Pred. No. 5.8e-267;
Matches 1622; Conservative 0; Mismatches 798; Indels 38; Gaps 3;
QY 182 AAAATTAAACAAGAGAGTAATAGGTGAAACTCACACCCAAAGAGCAAGAAAGTTCTTGG 241
Db 78 ACACITTAGAATAGGAGATGAGATGAACTCACCCCAAGAGATTAGATTAAGTTGATG 137

QY 242 TTATATTATCGGCGCGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTTAAAGCTCAACCAA 301
Db 138 CTCCTACTCGCTCGAGAAATTTGGCTAAAAAAGCGCAAGAAAGGCAATTAAGCTTAACCTAT 197
QY 302 CCGAAGACCATTCCTTACATTAGTCCCATATTATGACAGAGCGCGCTGGAAAAAAA 361
Db 198 GTAGAAGCAGTAGCTTTGATTAGTCCCATATTATGGAAGAGCGAGAGCTGGTAAAG 257
QY 362 ACCGTTGCCAGCTTATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGTAAATGCC 421
Db 258 ACTGCGCTGAATTGATGCAAGAGGCGCACTCTTTAAAAAAGATGATGATGAT 317
QY 422 GGGGTGGGTAAATATGGTTCCCGATCTAGGTGTGAGCCACCTTTCTCTGATGGTACGAAA 481
Db 318 GCGTGGCAAGCATGATCCATGAAGTGGGTATTGAAGCGATGTTTCTCTGAGGACTAAA 377
QY 482 CTTGTAACTGTGAATTTGGCCCATCGAACAGATGAGCACTTCAAGCGGGGGAAGTGA 541
Db 378 CTGTTAACCGTGCATATACCCCTATTGAGGCCAATGGTAAATTTAGTTCTCTGGTGA 434
QY 542 TTTGGTTGCGATAAAGACATCGAGCTCAATGCGAGGCAAGAGTAACCGAACTTTGAGTT 601
Db 435 TTTCTAAAAAATGAAGACATCACTATCAAGAGGCAAGAAAGCGTTAGCGTGAAGTT 494
QY 602 ACTAATGAAGGGCCTAAATCTTTGCGATGTGGTAGCCATTTCCACCTCTTTTGAAGCTAAC 661
Db 495 AAAAATGTTGGCGACAGACGGTTCAATCGGCTCACACTTCCATTTCTTTGAAGTGAAT 554
QY 662 AAGGCACTAAATTCGATCGTGAAGAAAGCCTATGCGAAAGCGCTAGATATTTCCCTCTGGC 721
Db 555 AGATGCTTAGACTTTGACAGAGAAAAAACTTTCGGTAAACGGTTAGACATTTGCGAGGGG 614
QY 722 AACACGCTACGATTTGGGCGAGACAAACCGCAAGAGTGCAGTTGATTCCTCTGGTGGC 781
Db 615 ACAGCGGTAAAGATTTGAGCGCTGCGAAGAAAAATTCGTTAGAAATTTGATTTGCAITTTGGCGGT 674
QY 782 AGTAAAAAAGTGTATGCGCATGAACGGCTTGTGAAATAACATCGCGGATGAACGCCATAAA 841
Db 675 AACAGAAAGATCTTTGGATTTAAGCATTTGTTGATAGACAGCAGACAGCAAGCAAGCAA 734
QY 842 CATAAAGCGCTTGACAAAGCGAAATCTCACGGATTT----- 877
Db 735 AAAATTTGCTTTACACAGAGCTAAAGAGCGTGTGTTTCATGGCGCTAAAGCGATGACAA 794
QY 878 -----ATCAAGTAAAGGAGACTCCCATGAAA---ATGAAAAAACAAGATATGTAAAT 926
Db 795 TATGTAAAAACCAATTAAGGAGTAAAGAAATGAAAAAGATTAGCAGAAAAAGAAATATGTTCT 854
QY 927 ACCTACGGACCCACCAAGGCGATAAAGTCGCTTAGGAGATACCGATCTTTGGGCGAGAA 986
Db 855 ATGTATGGTCTTACTACAGCGATTAAGTGAATTTGGCGATACAGACTTGATCGCTGAA 914
QY 987 GTAGAACATGACTATACCACTATGCGAAGAACTTAAATTTGGCGGGGTAAACATATC 1046
Db 915 GTAGAACATGACTACACCACTTTATGCGAAGAGCTTAAATTCGTTGGCGGTAAACCCCTA 974
QY 1047 CGTAGGGGTATGGGTGAGAGCAATAGCCCTGATGAAAAACCCCTAGATTTTATGTCATCACT 1106
Db 975 AGAAGAGGCATGAGCCAATCTAACACCCCTAGCAAGAGAGTTGGATTTAAATTTATCACT 1034
QY 1107 AACGGCATGATTTACACTACCGGGATTTTACAAAGCCGACATTTGGGATTTAAAAACGGC 1166
Db 1035 AACGCTTTAATCGTGGATTACACCGGTATTTTAAAGCGGATTTTGTATTTAAAGATGGC 1094
QY 1167 AAAATCCATGGCATTTGGCAAGGAGGAAAAACAAGACATGCAAGATGCGTGAAGCCCTCAT 1226
Db 1095 AAAATCGCTGGCATTTGGTAAAGCGGTAAACAAGACATGCAAGATGCGTTAAAAACAAT 1154
QY 1227 ATGGTCTGGGTGGGCAAGAGCACTAGCAGGGGAAAGGTATGATTTATACCGCTGGG 1286
Db 1155 CTTAGCGTAGGTCTCTGCTACTGAAGCGCTTACCGCGTGAAGGTTTGTATCGTAAACGGCTGGT 1214

QY 1287 GGAATCGATTACACACCCACTTCCTTTCTCCACAACAATTCCTACCGCTCTAGCCAAAT 1346
Db 1215 GGTATTGACACACATCCACTTCATTTACCCCAACAATCCCTACAGCTTTTGAAGC 1274
QY 1347 GCGTTTACAAACATGTTTGGAGCGGCACACAGTCTGTAGATGSCACGAATGCGACTACT 1406
Db 1275 GGTGAACCAACATGATGTTGTTGGAAACCGGTCTGCTGATGGCACTTAATGGCACTACT 1334
QY 1407 ATCACTCCGGGCAAAATGGAATTTGCACCGCATGTTGCGCGCAGCAGAGAAGTAGTATTCATG 1466
Db 1335 ATCACTCCAGGCAGAGAAATTTAAATGGATGCTCAGAGCGGCTGAAGAATAATTCATG 1394
QY 1467 AATGTGGCTTTTGGGCAAGCAATAGCTCTAGCAAAAAACAACTTTGTAGAACAAAGTA 1526
Db 1395 AATTTAGTTCCTTGGCTAAGGTAACGCTTTTAAACGATCGAGCTTTAGCCGATCAAAAT 1454
QY 1527 GAAGCGGCGCATGTTGTTTAAATGTCATGAAGACTGGGGCACAACACCAAGTGGATC 1586
Db 1455 GAAGCGGTGCGATGTCCTTTAAATTCAGAAAGACTGGGGCACCACTCCTTCTGCAATC 1514
QY 1587 GATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTGTGATCCACACCGATACA 1646
Db 1515 AATCATGCGTTAGATGTTGGGACAAATACGATGTGCAAGTGTGATCCACACAGACACT 1574
QY 1647 GTCATGAGCGAGTTATGATGATGACACCCCTAAATGCAATGAACGGGCGCGCATCCAT 1706
Db 1575 TTGAATGAAGCGGTTGTGTAAGACACATATGGCTGCTATTTGTGACGCACTATGCAC 1634
QY 1707 GCTTACCAATTCAGGAGCGGTGGAGGACACTCACTCATGTTATCACCATGGCAGGC 1766
Db 1635 ACTTTCACACTGAAGCGCTGGCGGCGACGCTCCTGATATATTAAAGTAGCGGT 1694
QY 1767 GAGCTCAATATTTACCTCTCCACCACCCCACTATTCCTTATACCATTAATACCGTT 1826
Db 1695 GAACACAACTTCTCCGCTTCCACTAACCCCACTCCTTTCACCGTGAATACAGAA 1754
QY 1827 GCAGAACACTTAGACATGCTATGACATGCCACACCTAGACAAAACGATCCGCGAGAT 1886
Db 1755 GCAGAGCACATGGACATGCTTATGTTGTGCCACACCTTGGATAAAGCAATTAAGAAGAT 1814
QY 1887 TTCAATTTTCTCAAGCGGTATCGCCCGGCTCTATCCGGCTGAAGATGCTCCAT 1946
Db 1815 GTTCAGTTCGCTGATTCAGGATTCGCGCTCAAAACCAATTCGGCTGAAGACACTTTGCAT 1874
QY 1947 GATATGGGTGATTCGGATGACAAGCTCGGATTCGCAAGCAATGGCGGTGACAGGCGAA 2006
Db 1875 GACATGGGATTTTCTCAATCACAGTTCAGCTCTCAAGGATGGCGGTGGGTGAA 1934
QY 2007 GTCAATCTCTCGAACTTGGCAGACTGCGGATAAGAAATAAAGAAATTTGGTAAGCTTCCT 2066
Db 1935 GTTATCACTAGAACTTGGCAACAGCTGACAAAAACAAGAAATTTGGCGGCTTGAA 1994
QY 2067 GAAGATGGCAAGATACGATAATTTCCGATTAAGCGCTACATCTCCAATACACTATC 2126
Db 1995 GAAGAAAAAGCGGATAACGCAACTTCAGGATCAACCGCTACTTGTCTAAATACACCAAT 2054
QY 2127 AACCCCGCTTTGACCCACGCGTGAGCGAGTATATCGCTCTGTGGAAGAGGGCAAGATC 2186
Db 2055 AACCCAGCATCGCTCATGGATTAGCGAGTATGATGTTCACTAGAAAGTGGGCAAGTG 2114
QY 2187 GCGGACTTGGTGTGGAAATCTCGCTCTTTTGGCGTAAACCAACCAAAATCGTGATCAAA 2246
Db 2115 GCTGACTTGTGTTGAGTCCAGCATTTCTTTGGCGTGAACCCCAACATGATCATCAA 2174
QY 2247 GCGGATGTTGTTCTCTGAAATGGGCAATTTCTAAACGGCTGTGCGCCACTCCCAAA 2306
Db 2175 GCGGATTTCAATGCTTTAAGCCTAAGCGGATGCGAACCGCTTCTATCCCTTACCCCAAA 2234
QY 2307 CCGGTTTATACCGGAAATGTTTGGGCATCACCGCAAGCGAAATTTGACCCAGCATC 2366
Db 2235 CCGGTTTATACAGAAATGTTTCGCTCATCATGTTAAAGCTAAATACGATGCAAAATC 2294
QY 2367 ACTTTTGTTCCTCAAGTCGCCTATGAAATGGCGTGAAGAAAGAAAGCTGGGCTTAGAGCGC 2426

Db 2295 ACTTTTGTGTCTCAAGCGGCTTATGACAAAAGGCATTAAAGAAATTAGGACTTTGAAAGA 2354
QY 2427 CAAGTTCTACCGGTCAAAAACTGCGTAACATCACCAAGAAAGACTTCAAGTTCAAACGAC 2486
Db 2355 CAAGTGTTCGGGTAAAAAAATTTGAGAAATATCACTAAAAAAGACATGCAATTCACGAC 2414
QY 2487 AAAACGGCAAAAAATCACCGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTC 2546
Db 2415 ACTACTGCTCACATTTGAAGTCAATCTCTGAAACTTACCATGTGTTCTGTTGGTGGCAAGAA 2474
QY 2547 TGCACTCTAAACCCACTCGCAAGTCCCTCTAGCCAGCGCTACACTTTTCTCTAGG 2604
Db 2475 GTAACTTCTAAACAGCAATAAAGTAGTGCTGGCAACTCTTTAGCAATTTTCTAGG 2532

RESULT 3

US-08-467-822-19
; Sequence 19, Application US/08467822
; Patent No. 5843460
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,822
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno

```

; OTHER INFORMATION: sequence."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 756..759
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno
; OTHER INFORMATION: sequence."
US-08-467-822-19

Query Match      36.9%;   Score 1062.6;   DB 2;   Length 2619;
Best Local Similarity 65.5%;   Pred. No. 1e-263;
Matches 1606;   Conservative 0;   Mismatches 804;   Indels 41;   Gaps 2;

Qy 192 ACAAGGAGTAATAGTGAAACTCACACCCAAAGAGCAAGAAAAGTTCTTGTTATATTATG 251
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 29 ATRAGGAGTTTAGTAGGAAACTAACGCTTAAGAACTAGACAAGTTAATGCTCCATTATG 88
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 252 CGGCGAAGTGGCTAGAAAGCGCAAGACAGAGGCTTAAAGCTCAACCAACCCGAGCCCA 311
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 89 CGGCGAGATTGGCGAAGAACGCTTGGCGGTGTGTGAAGTCAATTCACCGAGCGG 148
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 312 TTGCTTACATTAGTGCCCATATTATGGACGAAGCGCGCTGGAAAAAACCCTTGCCTCC 371
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 149 TCGCGCTCATTAGCGGCGTGTGATGAAAGCGCGTGTGTAATAAAGCGTGGCGG 208
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 372 AGCTTATGGAAGAGTGTGATGCACTTTTGTGAAAAAAGATGAAGTAATGCCCGGGTGGTA 431
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 209 ATTTGATGCAAGAGGAGGAGCACTTGGCTTAAAAAAGAAAAATGTATGACGGGTAGCAA 268
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 432 ATATGGTTCCCGATCTAGGTGTAGAACCCACCTTCTCTGATGGTACGAACCTTGTAACTG 491
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 269 GCATGATTCATGAAGTGGGATTTGAAGCTAACTTCCCCGATGGAAACCAAGCTTGTAACTA 328
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 492 TGAATTTGGCCCATCGAAACCAAGTACGACATTCAAAGCGGCGCAAGTGAATTTGGTTGCG 551
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 329 TCCACACTCCGTTAGAGATAATGGCAAAATTAGCCCCCGGCGAGTCTTCTTAA---AA 385
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 552 ATAAAGACATCGAGCTCAATTCGAGGCAAGAGATPAACCGAACTTGAAGTTACTAATGAAG 611
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 386 ATGAGGACATTACTATTAAAGCCGCAAGAGCCATTAGCTTTGAAAGTGAATAAAG 445
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 612 GGCTAAATCCTTGATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAACAGGCCTAA 671
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 446 GCGATCGTGTGTCAGGTGGGATCACTTTCCAATCTTTCGAAGTGAATAAGCTCTTGG 505
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 672 AATTTCGATCGTGAATAAGCCTATGGCAAAACGCTTAGATATTCCCTCTGGCAACACGCTAC 731
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 506 ACTTTCGATCGGCMAAAGCTTTTGCMAACGCTTAGACATGCATCTGGAACACAGCGGTGC 565
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 732 GCATTTGGGGCAGGACAAACCCGCAAGTGCAGTTGATTCCTCTTTGGTGGCAGTAAAAAAG 791
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 566 GC'TTTGAAACCGGGAGGAAAAAAGTGTGAACTCATTGTGACATCGCGGGGAATAAGCGCA 625
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 792 TGATTTGCATGAACGGGCTTTGTGAATAACATCGCGGATGAACGCCATAAACAATAAGCGC 851
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 626 TCTATGGCTTTAA'TCTTTTGGTGGATCGCCAAGCCGATGCGGTGTGTAATAAATAATCGGCT 685
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 852 TTGACAAGGCGAAATCTCAGCGATTT----- 877
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 686 TAAAAACGGCTTAAGAAAAAGTTTTTGGGTCTGTAAACTCGGGTTGTGAAGCGACTAAAG 745
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 878 ----ATCAAGTAAGGAGACTCCCATGAAAAATGAAAAAACAAGAAATATGTAATACCTTACG 933
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 746 ATAAAACAATAAGGAAAAAACCATGAAAAAGATTTACGAAAAGAATAATGTTTCTATGTATG 805
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 934 GACCACCAAGGGGATAAAGTGGCTTAGAGATACCGATCTTTTGGGCGAGAGTAGAAC 993
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 806 GTCCCACTACCGGGGATCGTGTGTAGACTCGCGCACACTGATTTGATCTTAGAAGTGGAGC 865
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 994 ATGACTATACCTATGCGGCAAGACTTAATTTTGGCGGGGTAAACCTATCCGTGAGG 1053
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 866 ATGATTGCACCACTATGTGTGAAGAGATCAAA'TTTGGGGCGGTAAACCTATCCGTGATG 925
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 1054 GTATGGGTACAGACCAATAGCCCTGATGAAAAACCCCTTAGATTTTAGTCATCACTAAACGCGA 1113

```


Db 2006 GGATCCGCATGGGATTTCTGACTATGTGGGCTCTGTGGAAGTGGCAATATAGCCGACC 2065
Qy 2194 TGGTGTGTGGATCCTGCTCTTTTGGGCTAAACCCAAAATCGTGATCAAGGGCGGTA 2253
Db 2066 TCGTGTCTTTGGAGTCCGGCTTTCTTTGGCATTAAGCCCAATATGATTATTAAGGGCGGAT 2125
Qy 2254 TGGTGTCTTTCTGAAATGGGGGATTTCTAAACGCGTCTGTGCCACTCTCCCAACCGGTTT 2313
Db 2126 TTATTGCGCTCTCTCAATGGGGGATGCCAATGCGTCTATTCCCAACCCCTCAGCCGGTCT 2185
Qy 2314 ATTACCGCGAAATGTTTGGGCATCGCGCAAGGCGAAATTTGACACGAGCATCACTTTTG 2373
Db 2186 ATTACCGTGAATGTTTGGACACCATGCGGAACCAATTCGACCAATATCARTTCG 2245
Qy 2374 TTTCCAAAGTCGCTATGAAAATGGGCTGAAAGAAAGCTGGGCTTTAGAGCGCAAGTTC 2433
Db 2246 TGTCCCAAGCGGCTTACAGGCAGGGATCAAGAGAACTAGGCTAGATCGCGCGCAC 2305
Qy 2434 TACCGGTCAAAAATCTGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGACAAACGG 2493
Db 2306 CGCCAGTGA AAAACTGTGCGCAATATCACTAAAGAGGACCTCAAATTTCAACGATGTGACCG 2365
Qy 2494 CAAAAATCACCGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGCAAACTCTGCACCT 2553
Db 2366 CACATATTGATGCAACCCCTGAAACCTATAAGGTGAAGTGGATGGCAAGAGGTAACT 2425
Qy 2554 CTAAACCCACCTCGCAAGTCCTCTAGCCAGCGCTACACTTTCTCTAGG 2604
Db 2426 CTAAAGCAGCAGATGAATTGAGCTAGCGCAACTTTTATAATTTGTTCTAGG 2476

RESULT 4

US-08-432-697-19
; Sequence 19, Application US/08432697
; Patent No. 6248330
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Parabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid

; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno
; OTHER INFORMATION: sequence."
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 756..759
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno
; OTHER INFORMATION: sequence."
; US-08-432-697-19
Query Match 36.9%; Score 1062.6; DB 3; Length 2619;
Best Local Similarity 65.5%; Pred. No. 1e-263;
Matches 1606; Conservative 0; Mismatches 804; Indels 41; Gaps 2;
Qy 192 ACAAGAGTAATAGGTGAAACTCACACCCAAAGAGCAAGAAAGTTCTTGTATATATATG 251
Db 29 ATAAGGAGTTTAGGATGAAACTTAACGCTTAAGAACTAGACAAGTTAATGCTCCTATATG 88
Qy 252 CGGGCGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGTCAACCAACCCGAAGCCA 311
Db 89 CGGGCAGATTGGCAGAGAAACGCTTTGGCGGTGTGTGAACACTCAATTACACCGAAGCGG 148
Qy 312 TTGCTTACATTAGTGGCCCATATTATGACGAAAGCGCGCCGCTGGGAAAAACCCGTTGCC 371
Db 149 TCGGCTCATTAGCGGGCTGTGATGGAAGAGCGCGTGATGGTAATAAAGCGTGGCGG 208
Qy 372 AGCTTATGGAAGATGCATGCACCTTTTGTAAAAAGATGAAGTAATGCCCGGGTGGGTA 431
Db 209 ATTTGATGCAAGAGGCGAGGACTTGGCTTAAAAAGAAAAATGTGATGACGCGGTAGCAA 268
Qy 432 ATATGTTCCCGATCTAGGTGTAGAGCCACTTTCTCTGATGGTACGAACTTGTAACTG 491
Db 269 GCATGATTTCATGAAGTGGGATTTGAAGCTAACTCCCGGATGGAAACCAAGCTTGTAACTA 328
Qy 492 TGAATTGGCCCATCGAACACAGATGAGCACTTCAAAGCGGGCGAAAGTGAATTTGGTTGCG 551
Db 329 TCCACTCTCGGTAGAGGTAATGCGCAATTTAGCCCGCGGAGGTCTTCTTAAA--AA 385
Qy 552 ATAAAGCATCGAGCTCAATGCGAGCAAGAAAGTAACCGAACTTGAGTTTACTAATGAAG 611
Db 386 ATAGGACATTACTATTAAACGCGCAAGAAAGCACTTAGCTTGAAAGTGAATAAATAAG 445
Qy 612 GGCTTAAATCCTTCATGTGGGTAGCCATTTCCACTTCTTTGAAGCTTAACAGGCACATAA 671
Db 446 GCGATCGTCTGTGAGGTGGGATCACATTTCCACTTCTTCGAAAGTGAATTAAGCTCTTGG 505
Qy 672 AATTTCGATCGTGA AAAAGCCCTATGCAAAACGCTTAGATATTCCCTCTCGCAACACGCTAC 731
Db 506 ACTTCGATCGCGCAAAAAGCTTTTGCAAAACGCTTAGACATCGCGGGAATAAGCGCA 565
Qy 732 GCATTGGGGCAGGACAAACCCGCAAGTGCAGTTGATTCTCTTTGGTGGCAGTAAAAAAG 791
Db 566 GCTTTGAACCGGGAGGAAAAAGTGTGGAACCTCATTTGACATCGCGGGAATAAGCGCA 625
Qy 792 TGATTGGCATGAACGGGCTTGTGAATAACATCGCGGATGAACGCCATTAACATAAAGCGC 851
Db 626 TCTATGGCTTTAAATTCTTTTGGTGGATCGCAAGCGGATGCGGATGTTAAATAAAGCTCGGT 685
Qy 852 TTGCAAGGCGAAATCTCAGCGATT-----
Db 686 TAAAAACGCGCTAAAAAGAAAAAGGTTTGGGTCTGTAAACATCGCGTTGTGAAGCAGCTAAAG 745
Qy 878 ----ATCAAGTAAGGAGAGCTCCCATGAAAAATGAAAAAACAAGAAATATGTAATACCTACG 933
Db 746 ATAAACAATAGGAAAAACCATGAAAAAGATTTTACGAAAAAGAAATATGTTCTATGATG 805
Qy 934 GACCCACCAAGGGCGATAAAGTGCCTTTAGAGATACCGATCTTTGGGCGAGAGTAGAAC 993

Db 806 GTCCCACTACCGGGATCGTGTAGACTCGCGACACTGATTTGATCTTAGAAGTGAGC 865
Qy 994 ATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGGGTAAACACTATCCGTGAGG 1053
Db 866 ATGATTGCACCACTTATGTGTGAAGAGATCAAAATTTGGGGCGGTAAACACTATCCGTGATG 925
Qy 1054 GTATGGGTGACAGCAATAGCCCTGTATGAAAAACACCCCTAGATTTAGTCATCACTAAACGGA 1113
Db 926 GGATGAGTCAACCAATAGCCCTAGCTCTTATGAATTTAGATTTGGTGTCTCACTAACGCC 985
Qy 1114 TGATTATCGACTACACCGGATTTACAAAGCCGACATTTGGGATTTAAAAAGCGCAAAATCC 1173
Db 986 TCATTGTGGACTATACCGGCAATTTACAAAGCCGACATTTGGGATTTAAAGACGGCAAGATTG 1045
Qy 1174 ATGGCAATTTGGCAAGCGGAGAAACAAAGACATGCAAGATGGCGTAAAGCCCTCATATGGTCG 1233
Db 1046 CAGGCATTTGGCAAGCGGCAATTAAGGACATGCAAGATGGCGTAGATAAATCTTTGCG 1105
Qy 1234 TGGGTGTGGGCA CAGAAGCACTAGCAGGGAAGGTATGATTTATTCACGCTGGGGGAATCG 1293
Db 1106 TAGGTCTGTCTACAGAGGCTTTGGCAGCTGAGGCTTGATTTGTAACCGCTGGTGGCATCG 1165
Qy 1294 ATTCACACACCACTTCTTCTCCACAAATTCCTACCGCTCTAGCCAAATGGCGTTA 1353
Db 1166 ATACGCAATTTCACTTATCTCTCCCAACAAATCCCTACTGCTTTTGGCAGCGGGTTA 1225
Qy 1354 CAACCAATTTTGGAGCGGCGACAGGCTCTGTAGATGGCAGAAATGCACTTACTATCACTC 1413
Db 1226 CAACCAATTTTGGAGGCGACAGGACCTCGGATGGCAGAAATGCAACCACTCACTC 1285
Qy 1414 CGGGCAAAATGGAACTTGCAACCGCATGTGCGCGCAGCAGAAAGATTTCTATGAATGTGG 1473
Db 1286 CCGGACGCGCTAATCTAAAGAAATGTTGGTGTGCGCGAAGAAATACGCCATGAATCTAG 1345
Qy 1474 GCTTTTGGCAAGCAATAGCTCTAGCAAAAAACAACTTGTAGAACAAAGTAGAAGCGG 1533
Db 1346 GCTTTTGGCTAAGGGGAATGTGTCTTACGAACCTCTTTTACGCGCATCAGATTTGAAGCAG 1405
Qy 1534 GCGCGATTTGTTTAAATTCATGAAGACTGGGCGACCAACCAAGTGGCATGCATCACT 1593
Db 1406 GGGCGATTTGTTTAAATTCAGAGACTGGGGAAGCACCCTGCGAGCTATTCACCACT 1465
Qy 1594 GCTTGAGCTGGCAGATGAATCGATGTGCAAGTGTGTATCCACACCGCATACAGTCAATG 1653
Db 1466 GCCTCAATGTCGCGATGAATCGATGTGCAAGTGTGCTATCCACACCGATACCTTAAAG 1525
Qy 1654 AGCGAGGTATGTAGATGACACCTTAAATGCAATGAACGCGCGCATCCATGCTTACC 1713
Db 1526 AGCGGGCTGTGTAGAAGACACCTTAGAGCGATTTGCGGGCGCACCATCCATACCTTCC 1585
Qy 1714 ACATTGAGGAGCGGTGGGAGCACTCACTGATGTATCACCATGGCGAGGCGACTCA 1773
Db 1586 ACATGAAGGGGTGGGGTGGACACGCTCCAGATGTATCAAAATGGCAGGGGAATTTA 1645
Qy 1774 ATATTTCACTCCTCTCCACACCCCACTATTCCTTATACCAATTAATACGTTGCAAGAC 1833
Db 1646 ACATTTCACTCCTCTCTATCAACCCGACCACTTCTTCCACCAAAACATGAAGCGGAGC 1705
Qy 1834 ACTTACATGCTCATGACATGCCACCACTAGACAAACGCAATCCGCGAGGATTTACAAT 1893
Db 1706 ACATGACATGTTAATGGTGTGCCACCACTTGGATAAAGTATCAAGGAAGATGTGCACT 1765
Qy 1894 TTTCTCAAAGCGTATCCGCGCGGCTCTATCCGCGCTGAAGATGTGCTCCATGATAGG 1953
Db 1766 TTGCGGATTCGAGGATTCGCGCGGCAAACTATCCGCGCTGAAGACCAACTCCATGACATGG 1825
Qy 1954 GTGTGATCGGATGACAAGCTCGATTTGCAAGCAATGGGGCTGCGAGCGAAGTCAATC 2013
Db 1826 GGATCTTTTCTATCACCAGCTCGACTCTCAGGCTATGGACCGTAGGCGAGGTATCA 1885
Qy 2014 CTGCACTTTGGCAGACTGGGGTAAGAATAAAAAAGAAATTTGGTAGCTTCTCTGAAGATG 2073
Db 1886 CAGGCACTTTGGCAGACAGCAGACAAAAACAAAAAGAGTTTGGGCGCTTGAAGAGAAA 1945

Qy 2074 GCAAAGATAACGATAATTTCCGCAATTAAAGCGCTACATCTCCAAATACACTATCAACCCCG 2133
Db 1946 AAGCGGATAACGACAACTTCCGCAATCAAGCGCTACATCTCTAAATACACCATCAACCCCG 2005
Qy 2134 CTTTGACCCAGCGGTGAGCGAGTATATCGCTCTGTGGAAGAGGGAAGATCGCGACT 2193
Db 2006 GGAATCGCGCATGGGATTTCTGACTATGTGGCTCTGTGGAAGTGGGCAAAATACGCGGACC 2065
Qy 2194 TGGTGTGTGGAATCTTCCTGCTTTTGGCGGTAAACCCAAATCGTGATCAAAAGCGGTA 2253
Db 2066 TCGTGTCTTGGAGTCCGCTTTCTTTGGCATTAAGCCCAATATGATTTAAGGGCGGAT 2125
Qy 2254 TGGTGTCTTCTCTGAAATCGCGAATCTAAACGCTCTGTGCCCACTCCCAACCGGTTT 2313
Db 2126 TTATTGGCTCTCTCAATGGCGATGCCAATGCTCTATTCCCACTCAGCCCGTCT 2185
Qy 2314 ATTACCGGAAATGTTTGGGCAATCAGCGAAGGCGAAATTTGACACAGCATCACTTTTG 2373
Db 2186 ATTACCGTGAATGTTTTGGACACCATGGGAAAAACAAATTCGACACCAATATCACTTTTCG 2245
Qy 2374 TTTCCAAAGTCGCTATGAAATGGCGTGAAGAAAAAGCTGGGCTTAGAGCGCCAAAGTTC 2433
Db 2246 TGTCCCAAGCGGCTTCAAGCGAGGATCAAGAGAACTAGGGCTAGATCGCGGCGAC 2305
Qy 2434 TACCGGTCAAAAACTGCGGTAAACATCAACCAAGAAAGACTTCAAGTTCAACGACAAAAACG 2493
Db 2306 CGCCAGTGAANAATCTGCGCAATATCACTAAAGAGACCTCAATTCACGATGTGACCG 2365
Qy 2494 CAAAAATCACGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACT 2553
Db 2366 CACATATGATGTCAACCTGAAACCTTATAAGGTGAAAGTGGATGGCAAGAGGTAACT 2425
Qy 2554 CTAAACCCACTCGCAAGTCCCTTAGCCAGCGCTACACTTTCTCTAGG 2604
Db 2426 CTAAGCAGCAGATGAATTAGCGCTAGCGCAACTTTTATTAATTTGTTAGG 2476

RESULT 5

US-08-466-248-19
; Sequence 19, Application US/08466248
; Patent No. 6258359
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,248
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:		
APPLICATION NUMBER:	US 08/432,697	
FILING DATE:	02-MAY-1995	
CLASSIFICATION:	435	
ATTORNEY/AGENT INFORMATION:		
NAME:	Meyers, Kenneth J.	
REGISTRATION NUMBER:	25,146	
REFERENCE/DOCKET NUMBER:	03495.0137-02000	
TELECOMMUNICATION INFORMATION:		
TELEPHONE:	(202) 408-4000	
TELEFAX:	(202) 408-4400	
INFORMATION FOR SEQ ID NO:	19:	
SEQUENCE CHARACTERISTICS:		
LENGTH:	2619 base pairs	
TYPE:	nucleic acid	
STRANDEDNESS:	double	
TOPOLOGY:	linear	
MOLECULE TYPE:	DNA (genomic)	
FEATURE:		
NAME/KEY:	misc_feature	
LOCATION:	31..36	
OTHER INFORMATION:	/standard_name= "Shine-Dalgarno	
OTHER INFORMATION:	sequence.	
FEATURE:		
NAME/KEY:	misc_feature	
LOCATION:	756..759	
OTHER INFORMATION:	/standard_name= "Shine-Dalgarno	
OTHER INFORMATION:	sequence.	
US-08-466-248-19		
Query Match		
Best Local Similarity 36.9%; Score 1062.6; DB 3; Length 2619;		
Matches 1606; Conservative 0; Mismatches 804; Indels 41; Gaps 2;		
Qy	192	ACAGGAGTAATAGGTGAACTCACCCCAAGAGCAAGAAAGTCTCTTTATATATG 251
Db	29	ATAAGGAGTTTGGAGTAACTAAGCGCTTAAAGAACTAGACAAAGTTTAACTCATATG 88
Qy	252	CGGCGAAGTGGCTAGAAAGCGCAAGCAGAGGCTTTAAAGCTCAACCAACCGAAGCCA 311
Db	89	CGGCGAGTTGGCAGAGAACGCTTGGCGGTGGTGAAGTCAATTAACCGAAGCGG 148
Qy	312	TTGCTTACATPAGTCCCATATTATGACGAGCGCGCGTGGAAAAAACCCTTGCC 371
Db	149	TCGCGCTCATTAGCGGGCGTGTGATGAAAGGCGCGTGTGTTAAATAAAGCGTGGCG 208
Qy	372	AGCTTATGGAAGAGTGATGACATTTTGAAGAAAGATGAAGTATGCCCGGGTGGTA 431
Db	209	ATTGTATGCAAGAGGCGAGGACTTGGCTTTAAAGAAAGAAATGTGATGCGCGGTAGCA 268
Qy	432	ATATGTTCCCGATCTAGTGTAGAGCCACCTTTCTGTATGGTACGAACTCTGTAAC 491
Db	269	GCAATGATCATGAGTGGGATTTGAAGCTAACTTCCCGATGGAACCAAGCTTGTAACTA 328
Qy	492	TGAATTTGGCCCATCGAACCGATGAGCATTCAAAGCGGCGGAAAGTGAATTTGGTTGCG 551
Db	329	TCCACACTCCGATAGAGATTAATGCAATTTAGCCCGCGGAGGTCTTCTTAA ---AA 385
Qy	552	ATAAGACATCGAGCTCAATGCGAGCAAGAAAGTAAACCGAACTCTGTAATGAAG 611
Db	386	ATGAGGACATTTACTATTAAACCGCGCAAGAGCCATTTAGCTTGAAGTGAAGAAATAAG 445
Qy	612	GGCTTAATCTTCGATGGGTAGCCATTTCCACTTTCTTTGAAGCTTAAACAGGCATTA 671
Db	446	CGGATCGTCTGTGAGGTGGGATCACATTTCCACTTCTTCGAAGTGAATAGTCTTTGG 505
Qy	672	AATTTCGATCGTGAAGAACCTTATGGCAACCGCTAGATTAATTCCTTGGCAACCGCTAC 731
Db	506	ACTTCGATCGCGCAAAAAGCTTTTGAACCGCTAGACATTTGCATCTGGAACAGCGGTG 565
Qy	732	GCAATGGGCGAGGACAAACCCGCAAAAGTGCAAGTTTCCTTTGGTGGCAGTAAAAAG 791
Db	566	GCTTTGAACCCGGGAGGAAAAAGTGTGGAACTCATTTGACATCGGCGGGAATAAGCGCA 625

```
QY 1834 ACTTAGACATGCTCATGACATGCCACCCTAGACAAACGCGATCCGCGAGGATTTACAAT 1893
Db 1706 ACATGACATGTTAATGGTGTGCCACCACTTGGATAAAGATATCAAGGAAGATGTGCAGT 1765
QY 1894 TTTCTCAAGCCGTATCCGCGCGGTCTTATCGCGGCTGAAGATGTGCTCCCATGATGG 1953
Db 1766 TTGCGGATTCGAGGATTTGCGCGCCCAAACTATCGCGCTGAAGACCAACTCCCATGACATGG 1825
QY 1954 GTGTGATCCGATGACAAGCTCGATTTCGAAGCAATGGCGGTGCGAGCGAAGTGAATTC 2013
Db 1826 GGATCTTTTCTATCACCAGCTCGACTCTCAGGCTATGGACGCGTAGGCGAGTGATCA 1885
QY 2014 CTCGAATCTGGCAGACTCGCGATAGAAATATAAAGAAATTTGTAAGCTTCTCTGAAGATG 2073
Db 1886 CAGCGACTTTGGCAGACAGCAGACAAACAAACAAAGAGTTTGGCGCTTGAAGAGAGAAA 1945
QY 2074 GCAAGATTAACATATTTCCGATTAAGCGGTACATCTCCAAATACACTATCAACCCCG 2133
Db 1946 AAGGCGATTAACGACAACTTTCCGATCAAAACGCTACATCTCTAAATACACCATCAACCCCG 2005
QY 2134 CTTTGAACCAACGCGGTGAGCGAGTATATCGGCTCTGTGAAGAGGCGAAGATCGCGACT 2193
Db 2006 GGATCGCGATCGGATTTCTGACTATGTGGGCTCTGTGAAGTGGCAATATACGCCGACC 2065
QY 2194 TGGTGTGTGGAATCTCGCTTTTGGCGTAAAAACCCAAATCTGTGATCAAGGCGGTA 2253
Db 2066 TGTGCTTTGGAGTCGGCTTTCTTTGGCATTAAGCCCAATATGATTTAAGGCGGAT 2125
QY 2254 TGGTGTCTTCTGAAATGGCGATTTAAACGCGTCTGTGCCACTCCCAACCGGTTT 2313
Db 2126 TTATTTGGCTCTCTCAAAATGGCGATGCGCAATGCGTCTATTCCACCCCTCAGCCGCTCT 2185
QY 2314 ATTACGGGAATGTTGGGATCATCGGATCAAGCGGCAAAATTCACACCATCATCTTTT 2373
Db 2186 ATTACCGGAAATGTTGGACACCATGGGAAACAAATTCGACACCAATATCATCTTCG 2245
QY 2374 TTTCCAAAGTCCTTATGAAATGGCGTGAAGAAAGTGGGCTTAGAGCGCAAGTTTC 2433
Db 2246 TGTCCNAGCGCTTACAGGCGAGGATCAAGAGAACTAGGCTAGATCGCGGCGAC 2305
QY 2434 TACCGGTCAAAACTCCGTAACATCACCAAGAAAGACTTCAAGTTCAACGCAAAACGG 2493
Db 2306 CGCCAGTGAAGAACTGTGCAATATCACTAAAGAGGACCTCAAAATTCACAGATGTGACCG 2365
QY 2494 GAAAATCACCGTCATCGGAAACCTTCGAGTCTTTGTAGATGCGCAACTCTGCACCT 2553
Db 2366 CACATATTGATGTAACCCCTGAAACCTATAAGGTGAAAGTGAAGTGAAGAGGTAACT 2425
QY 2554 CTAACCCACCTCGCAAGTGCCTCTAGCCGCGCTACACTTTTCTTCTAGG 2604
Db 2426 CTAAGCAGCAGATGAATGAGCCTTAGCGCAACTTTATTAATTTGTTCTAGG 2476
```

RESULT 6

```
US-09-431-705-1
; Sequence 1, Application US/09431705
; Patent No. 6585975
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Londono-Arcila, Patricia
; APPLICANT: Freeman, Donna
; TITLE OF INVENTION: Use of salmonella vectors for
; FILE REFERENCE: 06132/060001
; CURRENT APPLICATION NUMBER: US/09/431,705
; CURRENT FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4824
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: includes sequences from Helicobacter pylori,
; OTHER INFORMATION: Salmonella typhimurium, and Escherichia coli
; NAME/KEY: CDS
; LOCATION: (2)...(31)
; NAME/KEY: CDS
; LOCATION: (41)...(61)
; NAME/KEY: CDS
; LOCATION: (65)...(799)
; NAME/KEY: CDS
; LOCATION: (803)...(2512)
; NAME/KEY: CDS
; LOCATION: (2516)...(2692)
; NAME/KEY: CDS
; LOCATION: (2696)...(2896)
; NAME/KEY: CDS
; LOCATION: (2900)...(3322)
; NAME/KEY: CDS
; LOCATION: (3326)...(3385)
; NAME/KEY: CDS
; LOCATION: (3389)...(3406)
; NAME/KEY: CDS
; LOCATION: (3410)...(3466)
; NAME/KEY: CDS
; LOCATION: (3470)...(3598)
; NAME/KEY: CDS
; LOCATION: (3602)...(3661)
; NAME/KEY: CDS
; LOCATION: (3665)...(3697)
; NAME/KEY: CDS
; LOCATION: (3701)...(3769)
; NAME/KEY: CDS
; LOCATION: (3773)...(3817)
; NAME/KEY: CDS
; LOCATION: (3821)...(3844)
; NAME/KEY: CDS
; LOCATION: (3848)...(3889)
; US-09-431-705-1
```

Query Match 36.7%; Score 1058.6; DB 3; Length 4824;

Best Local Similarity 65.7%; Pred. No. 1.4e-262;

Matches 1606; Conservative 0; Mismatches 799; Indels 38; Gaps 3;

```
QY 197 GAGTAATAGTGTAAACTCACACCAAGAGCAAGAAAGTTCTTGTATATATTATGCGGCG 256
Db 77 GAGATCTCATGAACTCACCCCAAGAGTTAGATAGTTGATGCTCCACTACGCTGGA 136
QY 257 GAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGAAGCCATTGCT 316
Db 137 GAATTTGGCTAAAGAAACGCAAGAAAGGCAATTAAGCTTAACCTATGTAGAAGCAGTAGCT 196
QY 317 TACATTTAGTCCCATATTTATGACGAAGCGCGCGTGGAAAAAAGAAACCGTTGCCAGCTT 376
Db 197 TTGATTTAGTCCCATATTTATGGAAGAGCGAGAGCTGGTAAAAAGACTTGGCGGCTGAATTTG 256
QY 377 ATGGAAGAGTGCAATGCACATTTTAAAAAAGATGAAGTAATGCCCGGGTGGGTAATATG 436
Db 257 ATGCAAGAGGGCGCACTCTTTTAAAAACAGATATGATGATGATGCGTGGCGAAGCATG 316
QY 437 GTTCCCGATCTAGGTGTAGAAGCCACTTTTCTGTATGGTACGAAACTTTGTAACCTGTGAAT 496
Db 317 ATCCATGAAGTGGTATTGAAGCGATGTTTCTGTATGGGACTTAACTCGTAACCGTGCAT 376
QY 497 TGGCCCATCGAACCGATGAGCACTTCAAGCGCGGGAAGTGAATTTGTTGGTGAATAAA 556
Db 377 ACCCTTATTGAGGCCAATGGTAAATTTAGTTTCTGTGTGAGTTG---TTCTTAAAAAATCAA 433
QY 557 GACATCGAGCTCAATGTCAGGCAAGAGTAACCCGAACTTCAGGTTTACTTAATGAAGGCGCT 616
Db 434 GACATCACTATCAACGAGGCAAGAAAGCGGTTAGCGGTGAAGTTTAAAAATGTTGGCGAC 493
QY 617 AAATCTCTGATGTGGGTAGCCATTTTCCACTTCTTTTGAAGCTTAACCAAGGCACTAAAAATTC 676
Db 494 AGACCGGTTCAATCGGCTCACACTTCCACTTCTTTTGAAGTGAATAGATGCTTAGACTTT 553
```


Query Match	Best Local Similarity	Score	1058.6;	DB 3;	Length	4824;
Mismatches	Conservative	0;	Mismatches	269;	Indels	38; Gaps
197	GAGTAAATAGGTGAACTCACACCCAAAGAGCAAGAAAGTTCTTGTTATATATATGCGGC	256				
77	GAGATCTCATGAAATCTACCCAAAGAGTTAGATAAGTTGATGCTCCATCGCTGGA	136				
257	GAGTGGCTAGAAAGCGCAAGCAGAGGCTTTAAAGCTCAACCAACCGAAGCCATTTGCT	316				
137	GAAATGGCTAAAGCGCAAGCAAGAAAGGCAATTAAGCTTAATATGTAGAGCAGTAGCT	196				
317	TACATTAGTCCCATATATATGCAAGAGCGCCCTGGAAAAAAGCCGTTGCCAGCTT	376				
197	TTGATTAGTGGCCATATATATGCAAGAGCGAGCTGGTAAAGAGACTGCGGCTGAATTG	256				
377	ATGGAAGATGTCATGCACTTTTGAAGAAAGATGAAGTAATGCCCGGGTGGTAATATG	436				
257	ATGCAAGAGGGCGCACTCTTTTAAAGCCAGATGATGTAGTGGATGGCGTGGCAAGCATG	316				
437	GTTCGGATCTAGGTGTAGAACCCACCTTTCTCGATGTGTAGCAACTTTGTAATGTGAAT	496				
317	ATCCATGAGTGGTATTGAAGCGATTTCTCGATGGCACTAAACTCGTAACCGTGCAT	376				
497	TGGCCCATCGAACCGATAGACACTTCAAGCGGCGGAAGTGAATTTGGTTGGCATAAA	556				
377	ACCCCTATTGAGCCAAATGGTAAATTAGTTCTCGTGGTGAATTG---TTCTTAAAAAATGAA	433				
557	GACATGACCTCAATGCGGCAAGAGATTAACCGAATTTGAGTTTACTTAATGAAGGCGCT	616				
434	GACATCACTATCAAGCAAGGCAAAAGACCGTTAGCGTGAAGTTTAAATTTTGGCGAC	493				
617	AAATCTTCGATGTGGTAGCCATTTCCACTTTCTTTGAAGCTTAACAGGCGCACTAAATTC	676				
494	AGACCGTTCAATCGGCTCACACTTCCATTTCTTTGAAGTGAATAGATGCTAGACTTT	553				
677	GATCGTGAAGGCTATGGCAAGCGCTAGATATTCCTCTGGCAACACGCTACGATT	736				
554	GACAGAGAAAAAATTTGGTTAAAGCGTTTAGACATTTGGAGCGGACAGCGGTGAAGATT	613				

```
Qy 1782 CCTCTCCACACCCGCCACTATTCCTATATACCATTAATACGGTTCAGAACACTTAGAC 1841
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1694 CCGCTTCCACTAACCCACCATCCCTTCCCGTGAATACAGAAAGCAGAGCACATGGAC 1753
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1842 ATGCTCATGACATGCCACCACTAGACAAACGATCCGCGAGGATTTACAAATTTCTCAA 1901
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1754 ATGCTTATGGTGTGCCACCACTTGGATTAAGAAAGCATTAAGAAAGATGTTTCGCTGAT 1813
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1902 AGCGGTATCCGCCCGGCTCTATCGCGCTGAAGATGTCTCATATGATGGGTGTGATC 1961
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1814 TCAAGGATCGCCCTCAACACATTCGCGCTGAACACACTTTCATGACATGGGATTTTC 1873
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1962 GCATGACAGCTCGGATTCGCAAGCAATGGGGCGTCGAGCGAAAGTATCTTCGAAT 2021
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1874 TCAATCACAGTTCCTGACTCTCAAGCGATGGGCGTGTGGGTGAAGTTATCACTAGAACT 1933
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2022 TGGCAGACTCGGATAGATTAAGAAATTTGGTAAAGCTTCTGAAGATGGCAAGAT 2081
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1934 TGGCAACACAGCTGACAAAAACAAGAAAGATTTGGCGCTTGAAGAAAGAAAGCGAT 1993
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2082 AACGATAATTTCCGCTATTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACC 2141
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1994 AACGACAACTTCAGATCAACCGTACTTGTCTAAATACACCTTAACCCAGCGATCGCT 2053
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2142 CACGGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGATCGCCGACTTTGGTGGTG 2201
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2054 CATGGGATTAGCGAGTATGAGTTTCAGTAGAAGTGGGCAAGTGGCTGACTTGGTATTG 2113
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2202 TGGATCTCTGCTTTTGGGTTAAACCCCAAAATCTGATCAAGGCGGTATGGTGTG 2261
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2114 TGGAGTCCAGCACTTCTTTGGGTGAACCCCAACATGATCATCAAGCGGATTCATTGG 2173
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2262 TTCTCTCAAAATGGCGGATTTCAAGCGCTCTGTGCCACTCCCAACCGGTTTATACCGC 2321
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2174 TTAAGCCAAATGGCGGATGGAACGCTTCTATCCCTACCCCAACCGGTTTATACAGA 2233
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2322 GAAATGTTTGGGCATCAGCGCAAGCGGAAATTTGACACACGACATCACTTTTGTTCAAA 2381
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2234 GAAATGTTGCTCATCATGTTAAAGCTAAATACGATGCAACATCACTTTTGTCTCAA 2293
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2382 GTGCGCTATGAAATGGCGTGAAGAAAGCTGGGCTTTAGCGCCAAAGTTCTACCGGTC 2441
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2294 GCGGCTTATGCAAGGCAATTAAGAAAGATTTAGGACTTGAAGACAAAGTGTTCGCGGTA 2353
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2442 AAAAACTGCGTAACTACCAAGACATTCAGTTTCAACGACAAACGCGCAAAATC 2501
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2354 AAAAAATGCAAAATATCACTAAAAAAGACATGCAATTTCAACGACATCTACCGCTCACT 2413
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2502 ACCGTGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCC 2561
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2414 GAAGTCAATCTGAACTTACCATGTGTTCTGGATGGCAAGAAAGTAACTTCTAAACCA 2473
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2562 ACCTGCAAGTCTCTAGCCAGCGCTACACTTTCTTTCTAGG 2604
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2474 GCCAATAAAGTGAAGTGGCGCAACTCTTTAGCATTTTCTAGG 2516
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 8

```
US-07-732-242C-8
; Sequence 8, Application US/07732242C
; Patent No. 5298399
; GENERAL INFORMATION:
; APPLICANT: Uozumi, Takeshi; Masaki, Haruhiko;
; APPLICANT: Hidaka, Makoto; Nakamura, Akira;
; APPLICANT: Maeda, Michihisa; Yoneta, Yasuo
; TITLE OF INVENTION: Gene of Urease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Frishauf, Holtz, Goodman & Woodward, P.C.
; STREET: 600 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
```

```
; ZIP: 10016-2088
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 0.72mb
; OPERATING SYSTEM: MS DOS
; SOFTWARE: ASCII Form
; CURRENT APPLICATION DATA:
; FILING DATE: 19910718
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JPN 2-210178
; FILING DATE: 10-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Herbert
; REGISTRATION NUMBER: 17081
; REFERENCE/DOCKET NUMBER: 910532/HG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)972-1400
; TELEFAX: (212)370-1622
; TELEX: 236268
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6131 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-07-732-242C-8
```

```
Query Match 25.6%; Score 737.6; DB 2; Length 6131;
Best Local Similarity 58.3%; Pred. No. 1.1e-179;
Matches 1442; Conservative 0; Mismatches 969; Indels 63; Gaps 6;

Qy 185 ATTTAACACAGGAGTAATAGGTGAACTCACCCAAAGAGCAAGAAAGTTCTTGTTA 244
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 317 ACTTAAATAGGAGGTATTCATGAACTGACTTCACTGTAATGGAAGAGCTCATGATT 376
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 245 TATTATCGGCGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCC 304
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 377 GTAGTGGGCTGACTTGGCCCGCGTCTTAAGAGCGGGCTTAAATTAATTTATCCT 436
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 305 GAAGCCATTGCTTACATTTAGTCCCATATATATGACGAAGCGCGCTGGAAAAAACC 364
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 437 GAAGCTGTGCAATGATTACATATGAAGTCTGAGAGGGGCGCGGATGG---AAAAACG 493
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 365 GTTCCCGAGCTTATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGTAATGCCGG 424
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 494 GTAGCTCAGTTAATGCAATACGGTGCACACGATTCTTCAAAAAGAAAGATGTAATGGAAGG 553
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 425 GTGGTAAATATGGTTCCCGATCTAGGTAGAAAGCCACTTTCTCTGATGGTACGAACTT 484
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 554 GTTCCGGAATGATCCCGGATATTCAAAATTGAGGCAACTTTCTCTGATGGAACAAAGCTT 613
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 485 GTAACTGTGAATTTGGCCCATCGAACCATGAGCACTTCAAGCGGGCGGAAGTGAAATTT 544
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 614 GTACGGTTTCATGACCCGATCCGTTAAATGAGGAGACGTCATGATACCGAGGAGTA 673
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 545 GGTTCGGAATAAGACATCGAGCTCAATCGCAGCAAGAAAGTAACCGAACTTGAGTTACT 604
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 674 TGTATTAATAAAGAAACCTATTTTATGCAATCAAAATAAGCAGACGATCAAGATTCGCGT 733
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 605 AATGAA---CGGCTTAATCTTTCATGTGGTAGCCATTTTCCACTTCTTTTGAAGCTAA 660
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 734 GTTAAACCGGGCGGATCGACCTGTTTCAGTTGGTTCCCATTTTCATTTTTTTTGAAGTGA 793
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 661 CAAGGCACATAAAATTCGATCGTGAAGAAAGCCCTATGCGCAACGCCCTAGATATTTCCCTCTG 720
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 794 TCAATCGCTTCAATTTTCATCGTGAAGAAAGCATTTTGCATGCGTGTGAATATTTCCGCTG 853
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 721 CAACACGCTACGCAATTTGGGGAGGCAAAACCCGCAAGATGCAAGTTCTCTCTCTGCTG 780
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 854 AACGCGGTTTCGCTTCGAGCCCGGAGATGCGAAAGAGTAGAAATAAATTTTCCATTTTCAGG 913
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


Qy	781	CAGTAA	AAAAAGTGA	ATTGGCA	TGAACGGGCTTGTGA	ATAACATCGCGGATGA	ACGCCATAA	840
Db	914	TGAACG	CAAAAGTGTATGGTTTAA	ATAATATGTAACGA	-----	-----	948	
Qy	841	ACATAA	AGCGCTTGACA	AGGCGAAATCTC	CACGGATTTATCA	AGTAAAGGAGACTCC	CATGA	900
Db	949	--ATG	GATCAGTTGAAATGGG	AAAAAGAAATGAGTT	-----	-----	986	
Qy	901	AAATG	AAAAAACAG	AAATATGTAA	TACCTACCGACCCACCA	AGGCGGATAA	AGTGC	960
Db	987	CGATG	CTCGAAAGCA	TATGCGGATATGTTG	GACCAACTGTGCGGCA	CGCCATTCGTT	1046	
Qy	961	TAGAG	ATACCGATCTTTTGGG	CAGAGTAGAA	CATGACTATACCA	CTTATGGCGA	AGAAC	1020
Db	1047	TGGCAG	ATTTCAGAAATCTTTTAT	CGAAATTTGAA	AGGACTATACAA	CTATGCGAGATG	GAGG	1106
Qy	1021	TTAAAT	TTTGGCGGGTAAAA	ACTATCCGTGAGG	GTATGGGTACAGACAA	T---AGCCCTG	1077	
Db	1107	TAAAG	TTTGGCGGCGCAGG	TGATCCGAGATGGA	ATGGGGGAGACTCTTT	GGCGACAA	1166	
Qy	1078	ATGAAA	CACCCCTAGATTTT	AGTTCATCACTAAC	CGGATGATTATCGACTAC	ACCGGGATTT	1137	
Db	1167	GCGAT	GAAATCGTCGATCTCGTATTA	ACAATTCGATTA	TGTTGATTA	CACAGGTATTT	1226	
Qy	1138	ACAAG	CCGCACTTTGGGATTTAAAA	ACGGCAAAATCCA	TGGCATTTGGCAAG	CGAGGCAAA	1197	
Db	1227	ATAAG	CAGATATCGGCATAAAA	AGATGGAATGATTC	CTCATAGGAAAAAG	CGGGGAACC	1286	
Qy	1198	AGCAC	ATGCAAGATGGCGTAAG	CCCTCATATGTCGTGGGTGTGGG	CACAGAAAGCACTAG	1257		
Db	1287	CGTTG	TTAATGAGCGGGTCG	-----ATATG	TGTGATTTGGAGCAG	CACAGAAGTCA	1340	
Qy	1258	CAGGG	AAGGTATGATTAATTA	ACCGCTGGGGAAATCGATTC	CACACACCCACTTCC	TTTCTC	1317	
Db	1341	CCGCA	GAAGGATGATTTGTGAC	AGCGGAGGAATAGATGCTCATAT	TTCACTTTATTTG	TGCC	1400	
Qy	1318	CACAA	CAATTCCTACCGCTCTAG	CCNAATGGCGTTTACA	CCATGTTTGGAGGCGG	CACAG	1377	
Db	1401	CTCAG	CAAAATCGAAACCGCTCTTTG	CAATCGGGTGTGACACTATGAT	TGGCGGAGGAACAG	1460		
Qy	1378	GTCTG	TGATGGCACGAATGCG	ACTACTATCACTCGGGCAAAATG	GAACTTGCACCGCA	1437		
Db	1461	GACCG	CTACAGGCACAAATGCC	ACTACTTTGTATCACCGGGGCC	TTGGATATCCATCGTA	1520		
Qy	1438	TGTTG	CGCGCAGACAGAGATGATTT	CTATGAATGTGGCTTTT	TTGGCAAAAGGCAATAGCT	1497		
Db	1521	TGCTT	CAAGCAGCGGAAGAAATCCCG	GATAAATCTGGGCTTTTT	TAGGAAAGGGAAC	CTGTT	1580	
Qy	1498	CTAGC	AAAAAACAATTTGTAG	ACAAGTAGAAACGGGGCGGAT	TGGTTTTAAATTCGATG	1557		
Db	1581	CAGAT	GAGGCTCTCTTAAAG	GAACAAATTTGAACGGGAGCGGT	GGGATTTAAAGCTTCA	CAG	1640	
Qy	1558	AAGAC	TGGGCAACAACCA	AGTGCATCACTGCTTGAG	CGTGGCAGATGAA	TACG	1617	
Db	1641	AAGAT	TGGGATCGACGG	CGCGCTATTGATACATGTTT	GAAATGGCGGATCGATATG	1700		
Qy	1618	ATGTG	CAAGTTTGTATCCAC	CCGATACAGTCAATGAG	CGGATTTATGTAGATGA	CACCC	1677	
Db	1701	ATGTG	CAAGTAGCGATTCATAC	AGACACTTTAAATGAA	AGCGGATTTGTGCGAGTACTT	1760		
Qy	1678	TAAAT	GCAATGAACGGGCGGCCAT	CCATGCCCTACCA	TTTGAGGAGCGGGTGG	AGGAC	1737	
Db	1761	TGAAG	CCATAGACGCTCGAGT	GATTCATACATACAGA	AGGGGCTGGCGGGGAC	1820		
Qy	1738	ACTC	ACTGATGTTATCACC	ATGGCAGGCGAGCTCA	ATATTTCTACCCCTCCT	CCACCA	1797	
Db	1821	ATGCT	CCGATATTAATA	AGCGCGGCTTC	CGGATATTTTGGCTTCTT	CCACGATC	1880	
Qy	1798	CCA	CTATTCCTCATAC	CAATTAATACGGTTC	GAGAACACTTTAGACAT	GTCTCATGACAT	1857	
Db	1881	CAACT	CGACCTTACTATCA	TACTCAATCTTTGA	AGAGATTTTAGATATG	TATGTTATGTTG	1940	

```

RESULT 9
US-08-967-513-1
; Sequence 1, Application US/08967513
; Patent No. 5783436
; GENERAL INFORMATION:
; APPLICANT: Robert P. Hausinger
; TITLE OF INVENTION: Mutant Urease and Method
; TITLE OF INVENTION: of Use For Determination
; TITLE OF INVENTION: of Urea
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA

```

1958	Qy	ACACCTAGACAAACCGATCCGGCAGGATTTACAAATTTCTTCAAAGCGGTATCCGCCCGG	1917
1941	Db	ACCACCTTAGACGCTAAATATTTCCAGAGGATATTCGTTTTTCCCGATTACAGCATACGGAAG	2000
1918	Qy	GCTCTATCCGGCTGAAGATGTCTCCATGATATGGTGTGATCGCGATGACAAGCTCGG	1977
2001	Db	AGACCATCCGGCGGAAGATGTTTTACATGATTTTAGCGTTTTACGATGATTCGTCG	2060
1978	Qy	ATTTCGAAGCAATGGGGCGTGCAGGCGAAGTGATTCCTCGAACTTCGGCAGACTGCGGATA	2037
2061	Db	ATTACAGGCGATGGGGCGAGTAGAGGAAGTGATCATTCGTACGTGCGCAACCGCTGACA	2120
2038	Qy	AGATAAAAAAGAAATTTGGTAAGCTTCCTGAAGATGCGAAGATACGATAATTTCCGCA	2097
2121	Db	AGATGAAAAAGCAAGAGGGAAGTTACAGAAGACCAATGGTGTGGAGAGCAACTTTTCGTG	2180
2098	Qy	TTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACGGCGTGAGCGAGT	2157
2181	Db	TGAACGTTATATGCGCAATATACGATCAATCCGGCCATTGCTCATGTTATTCGGATT	2240
2158	Qy	ATATCGGCTCTGTGAAGAGGGCAAGATCCCGCACTTGGTGTGTGGAAATCTCTGCCTTTT	2217
2241	Db	ATGTGGGTTCTGTTGAAGTGGGAAAAATTAGCTGATTTAGTGTGTGGAAATCTCTGCTTTT	2300
2218	Qy	TTCCGCTAAACCCAAATTCGTGATCAAAAGGCGGTATGGTGTCTCTCGAATCGGGC	2277
2301	Db	TTGGTGTGAACCTGAACCTGGTCTTAAAGGAGGAATGATTCCTTACGACATATGGAG	2360
2278	Qy	ATTCTAACGCGTCTGTGCCACCTCCCCACCGGTTTTATTACCGCGAAATGTTTGGGCATC	2337
2361	Db	ATCCCAATGCCAGCATTCGACACCGCAGCGGTTTTATATCGTCCGATGTTTGCAGCGA	2420
2338	Qy	ACGGCAAGGCGAAATTTGACACAGCATCATCTTTGTGTTTCCAAAGTCGCTATGAAAAATG	2397
2421	Db	AAGGAGATGCCAATATCAAACTCTATCACCTTTGTTTTCGAAAGCAGCGTATGAAAAAG	2480
2398	Qy	CGGTGAAGAAAGCTGGGCTTAGAGCGCAAGTCTTACCGGTCAAAAACTGCCGTAAACA	2457
2481	Db	GCATTCATGAACAGTTGGGTTTGAGAAAAAGGTGAACCACTGCATGGAAATTCGAAT	2540
2458	Qy	TCACCAAGAAAGACTTCAAGTTTCAACGACAAAAACGGCAAAAAATCACCGTCGATCCGAAA	2517
2541	Db	TGACGAAAAAAGATTTAAATTTTGAACGATAAACCCCAAAATTTGACGTCGATCTCAGA	2600
2518	Qy	CCTTCGAGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACCTCGCAAGTGCCTC	2577
2601	Db	CATATGAAGTAAAGGTAGACGGTCAATTAAGTGAATGTGAACCGGCAGAAAACTCGTCCCTA	2660
2578	Qy	TAGCCAGCGCTACACTTTCTCTAGGCACAATGCCCCCTTTGGGGGCGAGGTTATTTTAG	2637
2661	Db	TGGCAACCGGTATTTCTTATTTTGAAGGTAGAAAAACATGATGGTTGAAAAAGTAGTCG	2720
2638	Qy	GAATCTTCATCAAA	2651
2721	Db	GAACATCAGACA	2734

Qy	375	TTATGGAAGAGTGCATGCACTTTTTCGAAAGAAAGATGAAGTATGCCCAGGGTGGGTAAATA	434
Db	197	TGATGAGGAAGAGCCCGTCACTGCTGACCCGCGAGCAGGTGATGAGAGGGCGTCCCGGAAA	256
Qy	435	TGCTTTCCGATCTAGGTGTAGAGCCACCTTTCTGATGGTACGAAACTTGTAACTGTGA	494
Db	257	TGATCCCGGATATCCAGGTGGAAGCCACTTCCCGACGGCTCGAAAGCTGGTCAACGTTT	316
Qy	495	ATTGGGCCATCGAAACAGATAGCACTTCAAGCGCGCGAAGTGAATTTGGTTGCGATA	554
Db	317	ACAACCGATTATCTGAGGTAGCGCCATGATCCCGGTGATATACGTTAAGCCCGGTC	376
Qy	555	AAGACATTCGAGCTCAATTCGACGGCAAGAAAGTAAACGAACTTTGAGGTACTTAATGAAGGC	614
Db	377	---AGATAGCCCTGAATACCGCGCGGCAACCTGTGCGTGGTCTGTGAGAACCAACGGCG	433
Qy	615	CTAAATCTTGATGTGGGTAGCCATTTCCACTCTTTTGAAGCTAACAGGCACTAAAT	674
Db	434	ATCGCGCGATTCAGGTGCGTTCCGACTACCAATTTCCCGGAGGTAAACCGGCGCTGAAGT	493
Qy	675	TCGATCGTGAAGAGCCTATGGCAACCGCTAGATATTCCTCTGGGCAACACAGCTACGCA	734
Db	494	TCGACGTCAGAGCGCGCGCTATCGCTGAATATCCGCGGGCAACCGCGGTACGCT	553
Qy	735	TTGGGCGAGCAAAAACCCGCAAAAGTCAGTTGATTCCTCTTGGTGGCAGTAAAAAAAGTGA	794
Db	554	TTGAAACCGCGCAGAAACGCGAGTGCAGCTGCTGGCTTTCGCGGTTCACCGCGCGTCT	613
Qy	795	TTGCGATGAACGGCTTTGGAATAAATATCCGCGATGAACGCAATAACATAAGCGCTTG	854
Db	614	TCGGCTTCCGCGCGAGGTCAATGGGCGCTCTGAGGTAAACGAT-----	657
Qy	855	ACAAGCGAATCTCAGCGATTATCAAGTAAGGAGACTCCCATGAAATGAAAAACAA	914
Db	658	-----GAGTATATTTTCAAGCCAG	676
Qy	915	GAATATGTAATACTACCGGACCCACCAAGGCGATAAAGTGGCTTTAGGAGATACCGAT	974
Db	677	GCCTATGCCGATATGTTTCGCGCCCAACCGTCGCGCAGAGTGGCGCTTGGCAGATACCGAG	736
Qy	975	CTTTGGCGAAGGTAGAACATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGCG	1034
Db	737	CTGTGGATCGAGGTGGAGGACGATTTGACCACTTACGCGGAGAGGTCAATTTCCGCGCG	796
Qy	1035	GGTAAACTATCCGTGAGGTATGGGTACAGCAATAGCCCTGATGAAACACCCCTAGAT	1094
Db	797	GGCAAGTGTATCCGCGACCGCATGGCGCAGGACAGATGCTGGCCGCGCACTGTGTGCAC	856
Qy	1095	TTAGTCATCACTAACGCGATGATTATTCGACTACACCGGGATTTTACAAGCCGCAATTTGGG	1154
Db	857	CTGGTGTCTACAAACGGTGTATCGTGCATCATCTGGGGGATCGTTAAGCCGATATCGGC	916
Qy	1155	ATTAATAACCGCAAAATCCATGGCATTTGGCAAGCGAGGAAACAGGACATGCAAGATGGC	1214
Db	917	GTGAAGACGCGCGGATCTTCGCCATTCGCAAGCGCGCAACCCGACATCCAGCCCCAAC	976
Qy	1215	GTAAGCCCTCATATGCTGTGGGTGTGGGCACAGAACCATAGCAGGGGAAGGTATGATT	1274
Db	977	GTACC-----ATPCCCATCGGCGCTGCGACGGAAGTGTATCGCGCGCGGAAGGAAAAATT	1030
Qy	1275	ATTACCGCTGGGGAATCGATTACACACCCACTTTCTCTCCACAACAATTTCCCTACC	1334
Db	1031	GTACCCCGCGCGGATTCGATACCAATTTTACCTTGGATCTGTCCGACGAGCGGGAAGAG	1090
Qy	1335	GCTCTAGCCAATGGCGTTACAAACATGTTTGGAGCGCGCACAGTCTCTAGTAGTGCACG	1394
Db	1091	GGCTGGTCTCTGGCGTGAACCAATGGTCCGCGCGGCGACCGGCCCGCGCGGGCACC	1150
Qy	1395	AATCGCACTACTATCATCTCCGGGCAATGGAACCTTGACCGCATGTTTGGCGCAGCAAGAA	1454
Db	1151	CATGCCACCACTGCAACCCGGGCGCGTGGTATATCTCAGCATGCTGCAAGCGCGCGAC	1210

Qy	1455	GAGTATTCTATGAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAACAAC	TT	1511
Db	1211	AGCCTCCGGTCAATATCGCCTGTCTGGGCAAGGGAAACGTTTCTCAGCGGATGCCCTG		1270
Qy	1515	GTAGAAACAAGTAGAAGCGGGCGGATTTGGTTTAAATTCGATGAAGACTGGGGCAACA		1574
Db	1271	CGGAGCAGGTGGCGGAGGGGTTATTGGCCTGAAGATCCAGAGGACTGGGGCGCAC		1330
Qy	1575	CCAAGTCCGATCGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGATC		1634
Db	1331	CCGGCGCGATGACTGTGCGTTAACCGTCGCGATGAAATGGAATCCAGGTCGCGCCTG		1390
Qy	1635	CACACCGATACAGTCAATGAGGCGAGTTATGTAGATGCACACCTAAATCGAATGAACGG		1694
Db	1391	CACAGGACACCCGTGAATCGGCTTTTGTGGAAGACACCCCTGCGCCCATCGCGCGG		1450
Qy	1695	CGCGCATCATCCCTTACACATTTGAGGAGCGGGTGGAGGACACTCACTGATGTATTG		1754
Db	1451	CGACCATCACACCTTCATACCGAAGGGGCGCGGGGCCATGCGCGGACATCATC		1510
Qy	1755	ACCATGGCAGGCGAGCTCAATATTCTACCTCTCCACCAACCCCACTATTCCTATACC		1814
Db	1511	ACCGCTCGGCCACCCGAACATTTTGGCGTGTCTCCACCAACCCAAACGCTGCCCTAC		1570
Qy	1815	ATTAAATACGGTTGCGAAGACCTTAGACATGCTCATGACATGCCACACCTAGACAAAC		1874
Db	1571	CTCAACACCATCATGAACATCTCGATATGCTGATGTCTGCCACCATCTGGACCCGG		1630
Qy	1875	ATCGCGAGGATTTACAAATTTCTCAAACCGGTATCCGCCCGGCTCTATCGCGGCTGA		1934
Db	1631	ATCGCGAGGACGTGGCTTTTGGCGAGTCGCGCATTTGCGCGGAACCATCGCTGGGAA		1690
Qy	1935	GATGTGCTCATGATATGGGTGTGATCGCGATGACAAGCTCGGATTCGGAATTCGAAG		1994
Db	1691	GACGTGTCACGATCTCGSGGCTTCTCGCTCCCTCTCGATTCGAGGCCATGCGC		1750
Qy	1995	CGTGCGGCGAAGTGAATCTCGAATTTGGCAGACTGCGGATGAAGTAATAAAGAAAT		2054
Db	1751	CGCGTGGGGAAGTGAATCTCCACCTGCGCAGTGGCGCATCGCATGAAGGTGCAGCG		1810
Qy	2055	GGTAAGCTTCTGAAGTGGCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCC		2114
Db	1811	GGAGCGCTGGCGAGGAGACCGGGATACGACAATTTCCGCGTGAAGCGCTACATCGCC		1870
Qy	2115	AAATACACTATCAACCCGCTTTTGACCCACCGCGTGAAGAGTATATTCGGCTCTGGA		2174
Db	1871	AAATACACATCAACCGGCGTGAACCAACGATCGCACGAACTCGGATCCATTGAG		1930
Qy	2175	GAGGGCAAGATCGCCGACTTTGGTGTGGAATCTCGCTTTTTTGGCGTAAACCCAAA		2234
Db	1931	GTGGTTAAGCTGGTACCTCGTGGTCTGGTCAACAGCCTTCTTCGCGGTGAACCGGCC		1990
Qy	2235	ATCGTGATCAAGCGGTATGGTGTCTTCTGTAATGGGCGATTTCAACGGCTGTG		2294
Db	1991	ACCGTGATCAAGCGCGCATGATCGCCCATCGCGCGATGGGCGATATCAATGCTCTATT		2050
Qy	2295	CCCACTCCCAACCGGTTTATTACCGCGAAATGTTTGGGCATCAACGCGAAGGCGAAAT		2354
Db	2051	CCGACCCCGGACCGGTGCACTACCGCCGATGTTTGGCGCGTGGGCGCGCCGCCAT		2110
Qy	2355	GACACGACATCACTTTTGTTCCTTCCAAAGTCGCTATGAAAATCGCGTGAAGAAAAGCTG		2414
Db	2111	CACGTCCGCTCACTTCTCTGTTCGAGGCGGGCAGCCAAATGGCGTTGCCGAGCGGCTG		2170
Qy	2415	GGCTTAGAGCGCCAGTTCTTACCGGTCAAAAATGCGGTAACTACCAAGAAGACTTC		2474
Db	2171	AACCTGCGAGCGGATCGCGGTGTGTGAAGGCTGCCGTACCGGTGCAGAAAGCCGACATG		2230
Qy	2475	AAGTTCAACGACAAAAACCGCAAAATCACCGTCGATCCGAAAACCTTTCGAGGCTTTGTA		2534
Db	2231	GTGCACAAACAGCTGTCAGCCTTAACATCACCGTGCAGCGCCAGACCTATGAGGTGCGG		2290
Qy	2535	GATGGCAAACTCTGCACCTCTAAACCCACCTCGCAAGTGCCTCTAGGCCACGCGCTACACT		2594

Db 563581 ||||| TTTTGTAAAGGCAACTGTTCAACCTTAGATCCTCTGGGTGAGCAAAATTGAAAGCGGGTG 563522

Qy 1537 CGATTGGTTTTAAATTGCTAAGAGACTGGGGCAACACCAAGTGGGATCGATCACTGCT 1596

Db 563521 CATTAGGTTTAAATCCAGAGACTGGGTGCAACGCTGCGGTGATTGATCTGCCT 563462

Qy 1597 TGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGG 1656

Db 563461 TAAAGGTAGCAGATGAAATGGATATTCAAGTGGCCATTACACAGACACGCTAAATGAAA 563402

Qy 1657 CAGTTATGTAGATGACACCTTAATGCAATGAAGCGGCGCCATCCATGCGCTACACA 1716

Db 563401 GTGCTTTTTGGAAGACACGATGAAAGCGATTGATGACGAGTCAATTCATCTTCATA 563342

Qy 1717 TTGAGGAGCGGGTGGAGGACACTCACCTGATGTATCACTGGCAGGCGAGTCAATA 1776

Db 563341 CGAGGGCGAGGTGGTGGTATGTCACCTGACATCAATTAAGCAGCGATGATTCACAAG 563282

Qy 1777 TTCTACCTCTCTCACACCCCACTATTCCTTATACCATTAATACGGTTGCGAACAAC 1836

Db 563281 TATTACCTGCTTCAACCAACCGCACTCGTCCGTTTACCAAAAACACCATTTGATGAACATT 563222

Qy 1837 TAGACATGCTCATGACATGCGCACCTAGACAAACGATCCGCGAGGATTTACAATTTT 1896

Db 563221 TGGATATGTTGATGGTTTGGCCATCACTTAGATAACGCGTCCGGAAGACGTAGCTTTTG 563162

Qy 1897 CTCAAGCCGTATCCGCGCGGCTCTATCCGCGTGAAGATGTCTCCATGATATGGGTG 1956

Db 563161 CGGATAGCGGTATCCGCGCTGAACCAATTCAGCAGAGATATTTTGCATGATATGGCG 563102

Qy 1957 TGATCGCGATGAACGCTCGGATTCGCAACCAATGGGGCGTGCAGCGCAAGTATCCCTC 2016

Db 563101 TCTTCTCAATTATGAGTTTCAGACTCTCAAGCGATGGGACGTATTGGCGAAGTCGTTATTC 563042

Qy 2017 GAATTGGCAGACTGGGATGAAGATTAAGATAAAAAGAAATTTGTAAGCTTCCTGGAATGCGA 2076

Db 563041 GTACATGGCAACTGCGAGATAAGATGAATAAATGCAACGTGGTAGCTAGGTAATGAAGGA - 562983

Qy 2077 AAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAATACACATCAACCCCGCTT 2136

Db 562982 -----AACGATACTTCGTTAAGCENATATTCGGGAATACACCATCAACCCAGCAA 562928

Qy 2137 TGACCCACGGCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCCGACTGG 2196

Db 562927 TTGCACATGTTATTCGCGAGCATATTGGCTCGTTAGAGTGGGTAAATTCGACATATCG 562868

Qy 2197 TGGTGTGAATCTGCGCTTTTTTGGCGTAAACCCCAAAATCGTGATCAAGCGCGGTATGG 2256

Db 562867 TGTTATGGAACACCGATGTTCTTTTGGCGTAAACCTGAAGTGGTGATTAATAAAGGCTTTA 562808

Qy 2257 TGGTCTTCTGAAATGGCGGATTTCTAACGCTGTGCGCCACTCCCGCAACCGGTTTATT 2316

Db 562807 TTAGCTATGCGAAATGGGCGATCCAAATGCTCAATTCCAACACCGCAACCTGTATTCT 562748

Qy 2317 ACCCGAAATGTTTGGGCACTACCGCAAGCGAAATTTGACACCGCATCACTTTTGTTT 2376

Db 562747 ACCGTCCAAATGTCAGGTGCAAGGCTTAGCAACCGCACAAACAGCAGTATTTCTGTTT 562688

Qy 2377 CCAAAGTGCCTATGAAAAATGGCGTGAAAGAAAGCTGGGCTTAGAGCGCCAAAGTTCTAC 2436

Db 562687 CACAAGCGCTGAAAAAGCTGATATTGTCGGAAGTTCGGTTTACACAAAGAAACCATTTG 562628

Qy 2437 CGGTCAAAACCTGCGTAACATCACCAAGAAAGACTTCAAGTTCAACGACAAACGGCAA 2496

Db 562627 CTGTGAAAGGTCGCGCAACGTAAGTAAAAAGATCTGGTTTCAATATGATGAACACCAA 562568

Qy 2497 AAATCACCGTGCATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACTCTA 2556

Db 562567 ACATTACTGTTGATGCTGAACGTTATGAAGTTCAGTTCGAGCGGAGAGATTAATTACCTGTG 562508

Qy 2557 AACCCACTCGCAAGTGCCTCTAGCCCGCGCTACACTTTCTTCTTA 2602

Db 562507 AACCACTGGATAGGTACCATTTGGGTCAAGCATATTTCTTCTTCTTA 562462

RESULT 12

US-09-643-990A-1/c

; Sequence 1, Application US/09643990A

; Patent No. 6528289

; GENERAL INFORMATION:

; APPLICANT: Robert D. Fleischmann

; Mark D. Adams

; Owen White

; Hamilton O. Smith

; J. Craig Venter

; TITLE OF INVENTION: the Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville,

; STATE: MD

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3 1/2 inch diskette

; COMPUTER: Dell Pentium

; OPERATING SYSTEM: MS DOS v6.22

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/643,990A

; FILING DATE: 23-Aug-2000

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/487,429

; FILING DATE: 1995-06-07

; APPLICATION NUMBER: 08/426,787

; FILING DATE: 1995-04-21

; ATTORNEY/AGENT INFORMATION:

; NAME: Kenley K. Hoover

; REGISTRATION NUMBER: 40,302

; REFERENCE/DOCKET NUMBER: PB186P1C1

; TELEPHONE: 301-610-5790

; TELEFAX: 310-309-8439

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1830121 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-643-990A-1

Query Match 23.0%; Score 664.4; DB 3; Length 1830121;

Best Local Similarity 57.1%; Pred. No. 1.1e-159;

Matches 1397; Conservative 1; Mismatches 1012; Indels 36; Gaps 9;

Qy 183 AAATTTAACAAGAGTATAGTGAACCTCACCCAAAGACGAGGGCTTAAAGCTCAACCAAC 242

Db 564897 ATATATAATTAAAGAAATGAAATATGCATTTAACTTCCAGAGAACAGAAAGTATGC 564838

Qy 243 TATATTATCGGGGAGTGGCTAGAAAGCCAAAGCAGAGGGCTTAAAGCTCAACCAAC 302

Db 564837 TTTTCTCGCGGCGAACTTGGCGCAAAACCGAAAGACGCGGCTTAAATTAACATATC 564778

Qy 303 CCGAAGCCATTGCTTACATTTAGTCCCCTATATTGGAAGACGCGCGTGGAAAAA 362

Db 564777 CAGAACTATTGCTTATATTGCTAGTCATTTACAAGAGGAGCAGC---AAGAGAAGGAATGA 564721

Qy 363 CCGTTGCCAGCTTATGGAAGAGTGCATGCACTTTTGAAGAAAGATGAAGTAAATGCCCG 422

Db 564720 GTGTAGCGGAAGTCAATGCAATATGCGCAACACTTTTAAACCGTTGATGATGTCATGGAAG 564661

QY	423	GGGTGGTAAATATGGTTCCGATCTAGGTGTAGAACCACTTTCTCTGATGGTACGAAAC	482	QY	1477	TTTTGGGCAAGGCAATAGCTCTAGCAAAAAAACAATTTGTAGAACCAAGTAGAAGCGGGCG	1536
Db	564660	GTGTGGCGGAATGGTTCATGAAGTCCAGATTGAAGCTACTTCCCGGATGGCAGCAAC	564601	Db	563581	TTTTGGTAAAGGCAACTGTTTCAACCCCTAGATCTCTCGTGTAGCAAAATTTGAAGCGGGTG	563522
QY	483	TTGTAACTGTGAATTCGGCCATC-----GAACAGATGAGACATTTCAAGCGGGCGGAAGTG	538	QY	1537	CGATTGGTTTTAAATTCGATGAAGACTGGGCAACAACCAAGTGGCATTCGATCAGTCTCT	1596
Db	564600	TTGTACCGTGCAATAATCAATCAGATAACCGWGGTGGGCTTTAGCCACAAAAATAA	564541	Db	563521	CATTAGGTTTTAAAAATCCAGAGACTGGGTGCAACGCTTGCCGTGATTGATTCTTGCT	563462
QY	539	AAATTTGGTTGGGATAAAGACATCGAGCTCAATGCGAGGCAAAAGATTAACCGAACTTGAG	598	QY	1597	TGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGGATACAGTCAATGAGG	1656
Db	564540	AATATCAATGGTGGCTAAAGCCACCCCTACAAGGAACAAGATATGATCCAGGGCAAT	564481	Db	563461	TAAAAGTAGCAGATGAATGATATTCAGTTGGCCATTACACAGACAGCTAAATGAAA	563402
QY	599	GTTACTAATCAAGGGCTTAAATCCTTGC--ATGTGGGTAGCCATTTCCCACTTCTTTGAAG	656	QY	1657	CAGTTATGTAGATGACACCCCTAAATGAATGAAGAGCGGGCGCCATCCATCGCTTACACA	1716
Db	564480	ACCAATTAGCCGAAGCGATATCTCGCTAAATGTCGCGAGAAAACCGTAATAATCGAAG	564421	Db	563401	GTGGCTTTTTGGAAGACACAGTGAAGCGATTGATGAGCAGAGTATTCATATCTTTCCATA	563342
QY	657	CTAACAGGCACTAAATTCGATCGTGTAAGAACCTATGGCAACGCCT-----AGATA	710	QY	1717	TTGAGGGAGCGGTGGAGGACACTCACCTGATGTATCACCATGGCAGCGAGCTCAATA	1776
Db	564420	TAAACAAATTCAGGCGACCGCCCAATTCAGTTGGCTCGCATTAACCATTTTTTTGAAACCA	564361	Db	563341	CGAGGGCGGAGGTGGTGGTCATGCACCTGACATCATTAAGACAGCGATGTATTCAAACG	563282
QY	711	TTCCCTCTGGCAACGCTACGCAATTTGGGCGAGACAAACCCGCAAGTGCAGTTGATTC	770	QY	1777	TTCTACCTCTCTCCACCAACCCCACTATTCCCTATACCATTAATACGGTTTCGAGAACACT	1836
Db	564360	ATAATGCCCTTAAATTTTGACCGCACCTTTGGCACGTGGAATGCGCTTAAATGTTCCATCTG	564301	Db	563281	TATTACTGCTTCAACCAACCCGACTCGTCGGTTTACCAAAAAACACCATTTGATGAACATT	563222
QY	771	CTCTTGGTGGCAGTAAATAAGTGATTGGGCATGAACGGGCTTGTGAATAACATCGCGGATG	830	QY	1837	TAGACATGTCATGACATGCCACCTAGACAAACGATCCGCGAGGATTTACAAATTTT	1896
Db	564300	GCAATGCGGTGGTTTTTGAAACCGGTGAAGTGAATCAGTGGAAATAGTTGCTTTTGGTG	564241	Db	563221	TGGATATGTTGATGGTTTGGCATCACCTAGATAAACGCTGCGGAGACAGCTAGCTTTTG	563162
QY	831	AACGCCATAAACAATAAGCGCTTTGACAAGCGGAAATCTCACGGATTTATCAAGTAAG--	888	QY	1897	CTCAAGCCGTATCCGCCCCGGCTCTATCCGGCTGAAGATGTGCTCCATGATATGGGTG	1956
Db	564240	GTAACCA--AATCATTTATGGTTTCCATAATCAAAATGATGGCAAAATTAAGGTAGGGCA	564182	Db	563161	CCGATAGCCGTATCCGCCCCGAAACCAATTCAGCAGAGAAGATATTTTGCATGATATGGCG	563102
QY	889	AGACTCCCATGAATGAATAAACAAGAATATGTAATACCTACGACCAACCAACCAAGGCG	948	QY	1957	TGATCCGATGACAAAGCTCGGATTCGCAAGCAATGGGCGGTGAGGCAAGTGTATTCCTC	2016
Db	564181	AGATGCGATTAACAATTTCAAGAGCGCAATATGTAGCAACTATGCTTCCAAACAGTTGGCG	564122	Db	563101	TCCTTCCATTATGATTTAGATTCAGACTCTCAGCGATGGACAGTATTTGGCGAAGTCTGTTATC	563042
QY	949	ATAAGTGGCTTTAGCAGATACCGATCTTTGGCGAGAAGTGAAGCAATGACTATPACCACCT	1008	QY	2017	GAATTTGGCAGACTGCGGATGAAGATAAAGAAATTTGGTAAGCTTCTCTGAAGATGGCA	2076
Db	564121	ATAAGTCGTTTGGCGATACCAATTTATGGCAACCAATGGCAAGATTTATTCAGCA	564082	Db	563041	GTACATGGCAAACTGCAGATAAGATGAATAATGCAACCTGGTGGCTAGGTAAATGAAGGA-	562983
QY	1009	ATGGCGAAGCACTTAAATTTGGCGGGGTAAACTATCCGTGAGGATTTGGGTGAGAGC-	1067	QY	2077	AGATTAACGATTAATTTCCGCAATTAAGCGCTACATCTCAAAATACACTATCAACCCCGCTT	2136
Db	564061	AAGGTGATGATGATAAATTTGGTGGCGGTAAAGCGTGGTGTGATGGCTCAAAGCG	564002	Db	562982	-----AACGATACTTCCGTATTAAACGATATATCGGGAATATACCACTAACCCAGCAA	562928
QY	1068	-----AATAGCCCTGATGAATAAACACCCCTAGATTAGTCATCACTAACGCGATGTTA	1119	QY	2137	TGACCCACGCGTGAAGGATATATCGGCTCTGTGGGAAGGGCAAGATGCCACTTGG	2196
Db	564001	GTACGCGCACTCGCGACAATCCGAATGTATTGGATTTTGTGATTAACCAACGTAATGATCA	563942	Db	562927	TTGCACATGGTATTTGGCGACATATTTGGCTGTTTAGAAGTGGTAAATTCAGATATCG	562868
QY	1120	TCGA---CTACACCGGATTTACAAAGCCGACATTTGGATTAATAACCGCAAAATCCATG	1176	QY	2197	TGGTGGGAATCTCGCTTTTTTTGGCGTAAAAACCAAAATCGTGATCAAAAGCGGTATGG	2256
Db	563941	TTGATGCTAAATTAAGGCAATTAATAAGCCGATATTTGGTATTCGTGATGGCGTATTTGGG	563882	Db	562867	TGTTATGGAAACCGATGTTCTTTGGCGTAAAACTGAACTGGTGTATTAATAAAGGCTTTA	562808
QY	1177	GCAATGGCAAGGCAAGAAACAAGACATGCAAGATGGCGTAAAGCCCTCATATGGTCGTGG	1236	QY	2257	TGCTCTTCTGAAATGGGCGATTTCTAACGCGTCTGTGCCACTCTCCCAACCGGTTTATTT	2316
Db	563881	GTAATGGACAAGAGTAAACCTGACACCATGATTAACGTACACCAATAATGATATCG	563822	Db	562807	TTAGCTATGCAAAATGGGCGATCCAAATGCTCAATTCCAACACCGCAACTGTATTCT	562748
QY	1237	GTGTGGGCAACAAGCACTAGCAGGGGAAGTATGATTAATACCGCTGGGGGAATCGATT	1296	QY	2317	ACCGCAAAATGTTTGGGCATCAGGCAAGGGAATTTGACACGAGCATCACTTTTGTGTTT	2376
Db	563821	GTGCAAGCAGGAAGTTTCAACCGTGCACATTTAATTTGCAACCGCTGGTGGTATCGATA	563762	Db	562747	ACCGTCAATGTATCGGTGCAAGGCTTAGCAACCGCACAAACAGCAGTATTCCTTTGTTT	562688
QY	1297	CACACCCCACTTCTCTTCCACACAATTTCCCTACCGCTCTAGCCATGGGTTACAA	1356	QY	2377	CCAAAGTCGCTATGAAATGGCGTGAAGAAAAGCTGGGCTTAGAGCGCCCAAGTCTTAC	2436
Db	563761	CCACATTTCACTTTATTTGTCCACAACAAGCAACAATGCAATGGAAGTGGGTTTACCA	563702	Db	562687	CACAAGCCGTGAAAAGCTGATATTCGTGGAGTTTCGGTTTACACAAGAAACCATTTG	562628
QY	1357	CCATGTTTGGAGCGGCACAGGTCTGTAGATGGCAGATGGACTACTATCACTCCCG	1416	QY	2437	CGGTCAAAAATCTCGGTAAACATCAACAAGAAAGCTTTCAAGTTCAACAGCAAAAACGGCAA	2496
Db	563701	CGTTAATTTGGTGGAACTGGCCCTGCTGATGGGTACACAGCAACCACTTTGATCCCTG	563642	Db	562627	CTGTGAAAGGCTCGCCCAACGCTAGGTAAAAAAGATCTGGTTTCATAATGATGAACACCAA	562568
QY	1417	GCAATGGAACTTGCAACCGCATTTTGGCGCAGCAGAGAAGTATTTCTATGAATTTGGGCT	1476	QY	2497	AAATCAACGTCGATCCGAAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACTCTA	2556
Db	563641	GCGCATGGTATATGGAACGTAATGTTTTCAAGCGCAGAAAGCCTTTCGCGGTAAACGTCGGAT	563582	Db	562567	ACATTACTGTGATGCTGAACGTTATGGAAGTTCAGTGGGAGGAGTTAATTAACCTGTG	562508
				QY	2557	AACCCACCTCGCAAGTGCCTCTAGCCAGCGCTACACTTTCTTCTTA	2602

[illegible]

QY 190 ACACAGGAGTAATAGGTGAAGTCAACCCAAAGAGCAAGAAAGTCTCTGTATATTA 249
Db 4028 AATTTGGGGTGTAAATTTTCATTTTACAAACGAGAGCAAGCAAAATTAATGATTAGT 3969
QY 250 TCGGGCGAAGTGGCTAGAAAGCGCAAGAGGCTTAAAGCTCAACCAACCCGAAGC 309
Db 3968 GCGGGCGAAGTGGCAACGTCGCTGAAGAGCAGCTGGTTTGAACATAAATCATCTCGAGC 3909
QY 310 CATTTGCTTACATTAGTCCCATATTATGACCAAGCGCGCTGGAAAAAACCCTTTGC 369
Db 3908 ATTAGCTTTAATCAGCGATGAATTATTAGAGGTGCACGCGATGG---TAAGACCGTTGC 3852
QY 370 CCAGCTTATGGAAGAGTGCATGCACTTTTGAAGAAAGATGAAGTAAATGCCGGGTGG 429
Db 3851 AGAGTTAATGAGTTATGGTAGACAAATTTCTAAACAAAGAGAGTGTATGATGGTGTGCA 3792
QY 430 TAAATATGGTTCCCGATCTAGGTGTAGAACGCCACCTTCTCTGATGGTACGAACTTGTAA 489
Db 3791 ACACATGATTACAGATATCGAAATCGAGGTACGTTCCCGATGGTACTAAGTTAATCAC 3732
QY 490 TGTGAATTTGGCCCATCGAACAGATGAGCACTTCAAAGCGGGC-GAAGTGAATAATTTGGTT 548
Db 3731 AGTACATCACCTATTGTTTAAAGGAGGAGTCAATATGATACAGAGGAAATTAATTACAA 3672
QY 549 GCGATAAAGACATCGAGCTCAATGCAAGGCAAGAAAGTAAACCGAACTTGAGGTTACTAA 608
Db 3671 AAAGTACAGAGGTGAAATTTAATAACCATCATCTCGAAACAGTTATCGAAGTTGAAATA 3612
QY 609 AAGGGCTTAATTCCTTGTGATGGGTAGGCATTTCCACTTTTGAAGCTAAACAGGCAC 668
Db 3611 CAGGAGATCGACCAATTTCAAGTGGGCTCACATTTTCAATTTTATGAAGCAAAATGCAAG 3552
QY 669 TAAATATCGATCGTGAAGAAAGCTATGGCAAGCCCTAGATATTCCTCTTGGTGGCAGTAAA 788
Db 3551 TAGATTTGCAAGTGAATATGCGATATGCGAATAATTTAGATATTCAGCTGGAGCAGCTG 3492
QY 729 TACGCAATTTGGGCGAGCAAAACCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAA 788
Db 3491 TTCGATTTGAACCTTGGGATTAAGAAAGTTCATTTAGTTGAAATGCTGGCAACGTA 3432
QY 789 AAGTGAATGGCATGAACCGGGCTTGTGAA---TAAACATCGCGGATGAACCCCATAAACATA 845
Db 3431 AAATTTTGGTTTTCGTGTATGGTCAATGGTCTATCGATGAGTCACGTCATCGCC 3372
QY 846 AAGCGCTGACAAAGCGGAATCTACCGATTTATC----- 880
Db 3371 CAACCTGATGAATAATGATGAATATGCAAGTGTATTCGGAGATAACGGTGTGAAAAACGTGA 3312
QY 881 -----AAGTAAGGAGACTCCCATGAAATGAAAAACAAAGAAATATGTAATAC 928
Db 3311 ATAAAAAGGAGGAAAAAGATCATGAGCTTTAAATGACGCAAAATCAATATACGAGCTT 3252
QY 929 CTACGGACCCCAAGGCGATAAAGTGGCTTTAGGAGATACCGATTCCTTTGGGCGAGAAGT 988
Db 3251 ATACGGTCAACTGTGTGGAGATTCCTATCGTTAGGTGATACGAATCTATTGCTCAAT 3192
QY 989 AGAACATGATATACCACTATGGGAGAACTTAAATTTGGCGGGGTAAACATATCCG 1048
Db 3191 AGAAAAAGACTATGCGGTTTATGGTGAAGAGCTACTTTTTGGTGGTGTGTAATCTATTAG 3132
QY 1049 TCAGGCTATGGGTACAGCAATAGCCC-----TGATGAAAAACACCTAGATTTAGT 1099
Db 3131 AGACGGTATGGCGCAAAATCCCTCGTGAACATGATGATGAGCTGTAACGTTGCAACCTTGT 3072
QY 1100 CATCTAACCGGATGATTATCGACTACACCGGGATTTTCAAGCGCGCATTTGGGATTTAA 1159
Db 3071 CATTTCTAATCGGTTATTATCGATTACGATAAAGTGGTTTAAAGCTGATATAGGCAATTAA 3012
QY 1160 AAACGCGAAATTCATGGCATTTGGCAAGCGCAGAAACAAAGACATGCAAGATGGCGTAAG 1219
Db 3011 AAATGGTTATATTTTCGCCATAGGTAAATGCCGCAACCCAGATATATATGGATAATGT--- 2955
QY 1220 CCTCATATGGTTCGTGGGTGGGCACAGAACACTAGCAGGGAAGGTATGATTATTAC 1279

Db 2954 ---CGCATTTATTATAGGTTCAACAAACAGATATCATTTGCCGCTGAAGGTAAAAATCGTCAC 2898
QY 1280 CGCTGGGGGAATCGATTTCACACACCCACTTCTCTCCCAACAAATTCCTCTACCGCTCT 1339
Db 2897 TGTGTGGTGTATTGATCTACTCATGTTCAATTTTAAATCTCGAACAGCAGAGGTGCGATT 2838
QY 1340 AGCAATTTGGGCTTACAAACCATTTTGGAGCGGCACAGGTCCTGTAGATGCAAGAAATGC 1399
Db 2837 AGAAAGTGGTATTACGACTCATATTGGTGGTGTACTGGTGTCTCAGAAGGTTCTAAAGC 2778
QY 1400 GACTACTATCATCTCGGGCAAAATGGAACCTTGCAACCGCATTTGCGCGCAGCAGAGAGTA 1459
Db 2777 AACAACTGTAACCTCCAGGTCATATGATGTAATTTTGAAGCTGCGGAAGGTTT 2718
QY 1460 TTTCTATGAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAAACAATTTGTAGA 1519
Db 2717 ACCGATTAATGTGCGTTTTACAGTTAAAGCAAGCAACAAATCCAACTGCACATCTTGA 2658
QY 1520 ACAAGTAGAAGCGGGCGCGATTGCTTTTAAATTTGCATGAAGACTGGGGCAACACCAAG 1579
Db 2657 ACAATCAATGCCGAGCAATTCGATTTAAAGTACATGAAGACTGGGGTGCACACCATC 2598
QY 1580 TGGGATCGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACAC 1639
Db 2597 TGTCTTGAAGTCAATGATGTTGCTGATGAATTTGATGTTCAAATTTGCAATTTACATGC 2538
QY 1640 CGATACAGTCAATGAGCGCAGGTTATGTAGATGACACCTAAATGCAATGAACGGCGCGC 1699
Db 2537 AGATATCTTTAAATGAAGCAGGATTTATGGAAGACACAATGGTGTCTGTTAAAGACCGTGT 2478
QY 1700 CATCCATGCTACCAATTTGAGGAGCGGGTGGAGGACACTCACCTGATGTTATCACCAT 1759
Db 2477 ACTTCATATGATACCATCTGGAAGTGTGTTGGCGGTCTATCGGCTGATTTAAATTAATC 2418
QY 1760 GGCAGCGAGCTCAATTTACCTCTCCCAACACCCCACTATTCCTATATACATTA 1819
Db 2417 CGCTGCAATTTCAAATATTTTACCTTCATCTACAAATCCAACTTTTGCCTTATACACATA 2358
QY 1820 TACGGTTGAGAACACTTAGACATGCTCATGACATGCGCACCACTAGACAAACCGCATCCG 1879
Db 2357 TACTGTAGATGAACATTTAGATATGGTAATGATTTACTCACCATTTTAAATGCGGCTATTCC 2298
QY 1880 CGAGGATTTACAAATTTTCTCAAAGCGGTATCCGCGCGGCTCTATCCGGCTGAAGATGT 1939
Db 2297 TGAAGATATCGCATTCGAGATTCAGTATTCGTTAAAGAAACGATTCGAGCAGAGATGT 2238
QY 1940 GCTCCATGATATGGGTGTGATCGGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGTGC 1999
Db 2237 TCTGCAAGATATGGGTGTATTTCAGTATGATTAGTTCCGATTCACAAGCAATGGGCGGTGT 2178
QY 2000 AGCGAAGTGTATTCCTCGAACTTGGCAGACTGGGATAGGATAAGAAATAAAGAAATTTGGTAA 2059
Db 2177 AGGTGAAGTAAATTTACACGAACATGCGCAAGTAGCACATCGCATGAAAGAACAACTGGTGC 2118
QY 2060 GCTTCTGAGATGGCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATA 2119
Db 2117 TTTAGATGGTGTATTTGAACATAATGATAATAATCGCATCAACGTTATATCGCTAAATA 2058
QY 2120 CACTATCAACCCCGCTTTGACCCACCGCGTGAAGAGTATATCGGCTCTGTGGAGAGGG 2179
Db 2057 TACAAATTAACCCAGCAATTTACATATGTTTCTGATATGTAGGATCTATCGAGCGGG 1998
QY 2180 CAAAGTCCGCGACTTGGTGTGGAAATCTCGCTTTTGGCGGTAAAAACCCAAATCGT 2239
Db 1997 CAACTAGCTGACATTTGCTTTATGGGACCCAAATTTCTTTGGGGTTAAACCTCGAATTAGT 1938
QY 2240 GATCAAAAGCGGTATGGTGTCTTCTGAAATGGGCGATTTAAACCGCTGTGGCCAC 2299
Db 1937 TGTAAAGGCGGATTAATTAACCTCTCGCGTAATGCGGATGCAAAATGGTTCTATACCTAC 1878
QY 2300 TCCCCAACCGGTTTATTACCGCGAAATGTTTGGGCAATCACGGCAAGCGAAATTTGACAC 2359

Db 1877 ATCTGAACCGATGAAGTACCGTAAATATGATGGTCAATAGCGGCGGAAACCTTACAAGTAC 1818
Qy 2360 CAGCATCACTTTTGTTCCTCAAGTCCGCTATGAAAAATGGCGTGAAGAAAGCTGGGCTT 2419
Db 1817 GTCAATGACATCGTCTCTAAACTGCTTATGAAATGGTATCAATCGTGCATTAATTT 1758
Qy 2420 AGAGCGCCAGTTCTACCGGTCAAAACTGCGGTAAACATCACCAGAAAGACTTCAAGTT 2479
Db 1757 AAAACGCATGGTGGCGTCCAGTTAAAAATATATAGACAAATATCTTAAAGCAGATATGAAAA 1698
Qy 2480 CAACGACAAAAACGGCAAAATACCCGTCGATCCGAAAAACCTTCGAGGTCTTTGTAGATGG 2539
Db 1697 TACAGTGCACACACCTTAAATGAGCGTGTATCCACAAACATATGAAGTATATAGATGG 1638
Qy 2540 CAAACTCTGCACCTCTTAAACCCACCTCGCAAGTGCCTCTAGCCAGCGCTTACATTTCTT 2599
Db 1637 AGAAAAATTAACAAGTAAATGCAGCAACTGAGTTACCATTAACCTCAAGATACTTCTTAT 1578
Qy 2600 CTAGGCACAA 2609
Db 1577 CTAGGAGGAA 1568

RESULT 15

US-08-781-986A-22/c
; Sequence 22, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-22

Query Match 23.0%; Score 662.8; DB 3; Length 5966;
Best Local Similarity 55.5%; Pred. No. 2.1e-160;
Matches 1437; Conservative 1; Mismatches 1093; Indels 59; Gaps 6;
Qy 70 TRCTATTATTAATTTTAAATAATTAATCTTATATCATATATAATAATTAATTAATTAATTA 129
Db 4148 TGCTTCAATATCGGATATGCTAAAAATATAAATTTTCTTAATTAACATTTGAACCTTTGTATA 4089

Qy 130 TTTAAAGTTAATAAAAGTAACGAAATTAGGACTATATAATCCCATTCGCTTTAAATTTA 189
Db 4088 TTAATTTGTTTATCACACCAATTAGGATTAATATAAATTTGTAAATTTGTATCTAATATGA 4029
Qy 190 ACACAGGAGTAATAGTGAACCTCACACCAAGAGCAAGAAAAGTTCTTTGTTATATTA 249
Db 4028 AATTGGGGTTTTAAATTTGCATTTTACAAACGAGAGACAAATTAATGATTTGTAGT 3969
Qy 250 TCGCGGCGAAGTGGCTAGAAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGGAAGC 309
Db 3968 GCGCGCGAAGTTGCACGCTCGTAAAGCAGCGTGTGAAACTTAAATCATCTCTGAGGC 3909
Qy 310 CATTTGCTTACATTTAGTGCCTATTTATGGAGAGCGCGCTGGAAAAAACAACCGTTGC 369
Db 3908 AATTAGCTTTAATCAGCGATGAATTTATTAGAAGTGCACGCGATGG---TAAGACCGTTGC 3852
Qy 370 CCAGCTTATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGTAAATGCCGCGGTGG 429
Db 3851 AGAGTTAATGAGTTTGGTAGACAAATCTTAACAAAGAGATGTCTGGATGGTGTGCA 3792
Qy 430 TAAATATGGTTCCCGATCTAGGTGTAGAAAGCCACTTTCTCTGATGGTACGAACTTTGTAAC 489
Db 3791 ACACATGATTACAGATATCGAATTCGAGGCTACGTTCCCGATGGTACTAAGTTAATCAC 3732
Qy 490 TGTGAATTTGCCCATCGAACCCAGATGAGCACTTCAAAGCGGC-GAAGTGAATTTGGTT 548
Db 3731 AGTACATCACCTTATTTTAAAGGAGGCGAGTCAATGATACCGAGGAAATTTATTACAA 3672
Qy 549 GCGATAAAGACATCGAGCTCAATGCGAGCAAGAAAGTAAACCGAACTTCAGGTTTACTAATG 608
Db 3671 AAAGTACAGAGTTGAAATTAATTAACCATCATCTCTGAAACAGTTATCGAGTTGAAATA 3612
Qy 609 AAGGCGCTTAAATCTTGCATGTGGTAGCCATTTCCACTTCTTTGAAGCTAACAAAGGCAC 668
Db 3611 CAGGAGATCGACCAATTCGAAGTGGCTCACATTTTCATTTTATGAAGCAAAATGCAGCAT 3552
Qy 669 TAAAAATTCGATCGTGAAGAAAGCCATATGCGAAAACGCTAGATATTTCCCTCTGCGCAACGC 728
Db 3551 TAGATTTGAAACGCTGAATGCGCATATGGAACACATTTAGATATTTCCAGCTCGAGCAGCTG 3492
Qy 729 TACGCAATTTGGGCGAGGACAAACCCGCAAGTGCAGTTGATTCCTCTTCGTCGCAATAAA 788
Db 3491 TTCGATTTGAACCTTGGGATAAAGAAAGTTCAATTAGTTGAATATCTCTGCAACGTA 3432
Qy 789 AAGTATTGGCATGAACGGGCTTGTGAA---TAACATCGCGGATGAACCCATTAACATA 845
Db 3431 AAATTTTGGTTTTCGTGGTATGTCATGTCCTATCGATGAGTCAGTGTCTATCGCC 3372
Qy 846 AAGCGTTGAACAGGCGAAATCTCAGGATTTATC----- 880
Db 3371 CAACTGATGAATAATGATGAATATGCAAGTGTATTTCGGAGATAACGGTGTGAAAAACGTGA 3312
Qy 881 -----AGTAAAGGAGACTCCCATGAATAAGAAAAACAAGATATGTAATAC 928
Db 3311 ATAAAAAAGGAGGAAAAAGATCATGAGCTTTTAAATGAACGCAAAATCAATATACGAGCTT 3252
Qy 929 CTACGACCCCAACCAAGCGGATAAAGTGCCTTAGGAGATACCGATCTTTGGGCGAAGT 988
Db 3251 ATACGCTCCAACTGTTGGAGATTCATTCGTTTAGGTGATACGAATCTATTGCTCAAT 3192
Qy 989 AGAATGACTATACCACTATGCGAAGAACTTAAATTTGCGCGGGTAAACATATCGG 1048
Db 3191 AGAAAAAGACTATTCGGGTTTATGTTGAAGAGCTACTTTTGGTGGTGAATCTATTAG 3132
Qy 1049 TGAGGGTATGGTTCAGAGCAATAGCCC-----TGATGAACAACCCATAGATTAGT 1099
Db 3131 AGACGGTATGGCGCAAAATCTCTGTTAAACACGTCGATGACGTCGAGCTTGCAGACTGT 3072
Qy 1100 CATCACTAAACCGCATGATTATTCGACTACACCGGGATTTACAAAGCGCAATTTGGGATTA 1159
Db 3071 CATTTCTAATGCCGTTATTTATCGATTACGATAAAGTGGTTAAAGCTGATATAGGCATTAA 3012
Qy 1160 AAACGGCAAAATCCATGGCATTTGGCAAGGCGAGGAAACAAGGACATGCAAGATGGCGTAAG 1219

[illegible]

Search completed: November 29, 2005, 07:30:10
Job time : 550 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2005, 00:08:23 ; Search time 2049 Seconds
(without alignments)
11635.245 Million cell updates/sec

Title: US-09-904-994B-1

Perfect score: 2883

Sequence: 1 rgragatttccarctt.....aaaaagtagagccacag 2883

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 413469005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA_Main:*

- 1: /cgn2_6/prodata1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/prodata1/pubpna/US09A_PUBCOMB.seq:*
- 4: /cgn2_6/prodata1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/prodata1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/prodata1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/prodata1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/prodata1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/prodata1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/prodata1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2880.6	99.9	2883	3	US-09-904-994B-1
2	2228	77.3	2452	3	US-09-904-994B-13
3	2158.6	74.9	2405	3	US-09-904-994B-4
4	2158.6	74.9	2407	3	US-09-904-994B-10
5	1948	67.6	2183	7	US-09-904-994B-7
6	1184.4	41.1	8407	7	US-10-639-273-1
7	940.8	32.6	1710	7	US-10-639-273-4
8	902.4	31.3	1719	8	US-10-476-313-10
9	886.4	30.7	1710	7	US-10-335-977-3849
10	885.8	30.7	1815	3	US-09-895-913A-251
11	884.8	30.7	1710	7	US-10-282-122A-22427
12	880	30.5	1710	9	US-10-500-447A-5
13	876.8	30.4	1717	8	US-10-476-313-11
14	855.6	29.7	2385	3	US-09-402-100-1
15	734.6	25.5	2341	7	US-10-282-122A-32950
16	690.6	24.0	2439	7	US-10-282-122A-34390
17	664.4	23.0	1830121	7	US-10-329-670-1
18	664.4	23.0	1830121	8	US-10-158-865-1
19	664.4	23.0	1830121	9	US-10-981-687-1
20	662.8	23.0	5966	2	US-08-781-986A-22
21	662.8	23.0	5966	7	US-10-329-624-22
22	658.6	22.8	1698	7	US-10-282-122A-33102
23	634.8	22.0	3919	9	US-10-795-159-460

24	634.8	22.0	85814	9	US-10-795-159-680
25	630.6	21.9	8729	5	US-10-114-170-258
26	630.6	21.9	87563	5	US-10-114-170-57
27	625.4	21.7	1719	7	US-10-282-122A-22026
28	617.2	21.4	1701	7	US-10-282-122A-23823
29	609.6	21.1	1704	7	US-10-282-122A-13582
30	603.4	20.9	1704	7	US-10-282-122A-11454
31	602.8	20.9	1719	7	US-10-282-122A-41843
32	591.4	20.5	1698	7	US-10-282-122A-8780
33	590.6	20.5	1704	7	US-10-282-122A-14054
34	587	20.4	1701	7	US-10-282-122A-32068
35	578.6	20.1	1722	7	US-10-724-972A-2754
36	577.4	20.0	1701	7	US-10-282-122A-30559
37	573.8	19.9	1707	7	US-10-282-122A-20639
38	573.8	19.9	1707	7	US-10-282-122A-20660
39	565	19.6	1713	7	US-10-282-122A-14888
40	556.6	19.3	1716	7	US-10-282-122A-8259
41	555	19.3	3309400	3	US-09-738-626-1
42	553	19.2	1710	3	US-09-738-626-99
43	548.2	19.0	2514	6	US-10-418-962-1
44	547.2	19.0	2688	7	US-10-437-963-21574
45	526	18.2	3110	7	US-10-424-599-111793

ALIGNMENTS

RESULT 1

US-09-904-994B-1
; Sequence 1, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoesesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2883
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (206)..(886)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (897)..(2603)
US-09-904-994B-1

Query Match 99.9%; Score 2880.6; DB 3; Length 2883;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	RGRAGATTTTCCARCACTTCAAGCACATATGATCTGTGTTGGTGTAAATTCR	60
Db	1	RGRAGATTTTCCARCACTTCAAGCACATATGATCTGTGTTGGTGTAAATTCR	60
Qy	61	ACTTGTTAATCTATTATTAATTTTAAATTAATTAATTAATTAATTAATTA	120
Db	61	ACTTGTTAATCTATTATTAATTTTAAATTAATTAATTAATTAATTAATTA	120
Qy	121	TTACTTATATAAAAGTTAATAAAGTAGTAAGAAATAGGACTATATCCATTC	180
Db	121	TTACTTATATAAAAGTTAATAAAGTAGTAAGAAATAGGACTATATCCATTC	180
Qy	181	TAAATTTAACAAGAGTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT	240
Db	181	TAAATTTAACAAGAGTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT	240
Qy	241	GTATATATTCGGCGGAAGTGGCTAGAAAGCGAAAGCGAGAGGGCTTAAAGCTCAACCA	300
Db	241	GTATATATTCGGCGGAAGTGGCTAGAAAGCGAAAGCGAGAGGGCTTAAAGCTCAACCA	300

Db 241 GTTATATTATCGGGCGAAGTGTCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCA 300
Qy 301 ACCGGAAGCCATTGCTTACATTAGTCCCAATTATGGACGAAGCGCGCGTGGAAAAA 360
Db 301 ACCGGAAGCCATTGCTTACATTAGTCCCAATTATGGACGAAGCGCGCGTGGAAAAA 360
Qy 361 AACCGTTGCCAGCTTATGGAAGAGTGCATGCACTTTTGGAAAAAGATGAATATGCC 420
Db 361 AACCGTTGCCAGCTTATGGAAGAGTGCATGCACTTTTGGAAAAAGATGAATATGCC 420
Qy 421 CGGGTGGGTGAATATGTTTCCCGATCTTAGTGTAGAAAGCCACCTTCTCTGATGTCGAA 480
Db 421 CGGGTGGGTGAATATGTTTCCCGATCTTAGTGTAGAAAGCCACCTTCTCTGATGTCGAA 480
Qy 481 ACTTGTAACCTGTGAATTTGGCCCATCGAACCCAGATGAGCACTTCAAAGCGGCGGAAGTAA 540
Db 481 ACTTGTAACCTGTGAATTTGGCCCATCGAACCCAGATGAGCACTTCAAAGCGGCGGAAGTAA 540
Qy 541 ATTTGGTTGGGATAAAGACATCGAGCTCAATGCAAGGCAAAAGATGAACCGAACTTGAGGT 600
Db 541 ATTTGGTTGGGATAAAGACATCGAGCTCAATGCAAGGCAAAAGATGAACCGAACTTGAGGT 600
Qy 601 TACTTAATGAAGGCGCTTAAATCCCTTTCATGTGGGTAGCCATTTCCACCTTCTTTGAAGTAA 660
Db 601 TACTTAATGAAGGCGCTTAAATCCCTTTCATGTGGGTAGCCATTTCCACCTTCTTTGAAGTAA 660
Qy 661 CAAGGCACTAAAAATTCGATCGTGAAGAAAGCCTATGGCAAGCGCTAGATATTCCTCTGG 720
Db 661 CAAGGCACTAAAAATTCGATCGTGAAGAAAGCCTATGGCAAGCGCTAGATATTCCTCTGG 720
Qy 721 CAACACGCTACGCAATTTGGGCGCAGGACAAACCCGCAAAAGTGCAGTTGATTCCTTTGGTGG 780
Db 721 CAACACGCTACGCAATTTGGGCGCAGGACAAACCCGCAAAAGTGCAGTTGATTCCTTTGGTGG 780
Qy 781 CAGTAAAAAGTCAATTTGGCATGAACCGGCTTGTGAATAACATCGCGATGAACGCCATAA 840
Db 781 CAGTAAAAAGTCAATTTGGCATGAACCGGCTTGTGAATAACATCGCGATGAACGCCATAA 840
Qy 841 ACATAAAGCGCTTGACAAGCGCAAAATCTCACGGATTTATCAAGTAAGGAGACTCCCATGA 900
Db 841 ACATAAAGCGCTTGACAAGCGCAAAATCTCACGGATTTATCAAGTAAGGAGACTCCCATGA 900
Qy 901 AAATGAAAAAACAAGAAATATGTAATACTTAACCTTAACCGACCCCAAGGCGATAAAGTGCCT 960
Db 901 AAATGAAAAAACAAGAAATATGTAATACTTAACCTTAACCGACCCCAAGGCGATAAAGTGCCT 960
Qy 961 TAGGAGATACCGATCTTTGGGCGAAGTAGAACAATGACTATACCACTATGGCGAAGAAC 1020
Db 961 TAGGAGATACCGATCTTTGGGCGAAGTAGAACAATGACTATACCACTATGGCGAAGAAC 1020
Qy 1021 TTAATTTGGCGGGTAAACCTATCCGTGAGGGTATGGGTACAGACAAATAGCCCTCATG 1080
Db 1021 TTAATTTGGCGGGTAAACCTATCCGTGAGGGTATGGGTACAGACAAATAGCCCTCATG 1080
Qy 1081 AAAACACCTTAGATTTAGTTCATCACTAACCGCATGATTTACGACTACACCGGATTTACA 1140
Db 1081 AAAACACCTTAGATTTAGTTCATCACTAACCGCATGATTTACGACTACACCGGATTTACA 1140
Qy 1141 AAGCGCATTTGGATTTAAAAAAGGCAAAATTCATGCGATTGGCAAGCGAGAAACAAAG 1200
Db 1141 AAGCGCATTTGGATTTAAAAAAGGCAAAATTCATGCGATTGGCAAGCGAGAAACAAAG 1200
Qy 1201 ACATGCAAGATGGGTAAAGCCCTCATATGTTGCGGGTGTGGGACAGAGACATAGCAG 1260
Db 1201 ACATGCAAGATGGGTAAAGCCCTCATATGTTGCGGGTGTGGGACAGAGACATAGCAG 1260
Qy 1261 GGGAAAGTATGATTTATACCGCTGGGGGAATTCGATTCACACACCCACTTCTTTCTCCAC 1320
Db 1261 GGGAAAGTATGATTTATACCGCTGGGGGAATTCGATTCACACACCCACTTCTTTCTCCAC 1320
Qy 1321 AACAAATTCCTTACCGCTCTAGCCCAATGGCGTTACAACATGTTTGGAGGCGGCAAGTTC 1380
Db 1321 AACAAATTCCTTACCGCTCTAGCCCAATGGCGTTACAACATGTTTGGAGGCGGCAAGTTC 1380

Qy 1381 CTGTAGATGCAAGATGCGAATCTACTATCACTCCGGGCAAAATGGAACCTTGCAACCCCATGT 1440
Db 1381 CTGTAGATGCAAGATGCGAATCTACTATCACTCCGGGCAAAATGGAACCTTGCAACCCCATGT 1440
Qy 1441 TGGCGCAGCAGAGAGATTTCTATGAATGTGGGCTTTTGGGCAAAAGCAATAGCTCTA 1500
Db 1441 TGGCGCAGCAGAGAGATTTCTATGAATGTGGGCTTTTGGGCAAAAGCAATAGCTCTA 1500
Qy 1501 GCAAAAAACAATTTGTAGAAACAAGTAGAAGCGGCGCGATTTGGTTTAAATTTGCAATGAAG 1560
Db 1501 GCAAAAAACAATTTGTAGAAACAAGTAGAAGCGGCGCGATTTGGTTTAAATTTGCAATGAAG 1560
Qy 1561 ACTGGGCGACAACACCAAGTGCATCACTCTCTTGAGCGTGCAGATGAATACGATG 1620
Db 1561 ACTGGGCGACAACACCAAGTGCATCACTCTCTTGAGCGTGCAGATGAATACGATG 1620
Qy 1621 TGCAGGTTGTATCCACACCGATACAGTCAATGAGGAGGTTATGTAGATGACACCCATAA 1680
Db 1621 TGCAGGTTGTATCCACACCGATACAGTCAATGAGGAGGTTATGTAGATGACACCCATAA 1680
Qy 1681 ATGCAATGAACCGGCGCGCATCCATGCTACCAATTTGAGGAGCGGCTGGAGCACT 1740
Db 1681 ATGCAATGAACCGGCGCGCATCCATGCTACCAATTTGAGGAGCGGCTGGAGCACT 1740
Qy 1741 CACTGATGTTATCACCATGCGAGCGAGTCAATATTTCTACCTCTCCACACCCCA 1800
Db 1741 CACTGATGTTATCACCATGCGAGCGAGTCAATATTTCTACCTCTCCACACCCCA 1800
Qy 1801 CTATTCCTATACCATTAATACGTTGCGAAGCACTTAGACATGCTCATGATGCCACC 1860
Db 1801 CTATTCCTATACCATTAATACGTTGCGAAGCACTTAGACATGCTCATGATGCCACC 1860
Qy 1861 ACCTAGACAAACGATCCGCGAGATTTACAATTTTCTCAAAGCGTATCCGCCCGGCT 1920
Db 1861 ACCTAGACAAACGATCCGCGAGATTTACAATTTTCTCAAAGCGTATCCGCCCGGCT 1920
Qy 1921 CTATCGCGCTGAAAGATGTGCTCATGATATGGGTGATCGCGATGACAAGCTCGGAT 1980
Db 1921 CTATCGCGCTGAAAGATGTGCTCATGATATGGGTGATCGCGATGACAAGCTCGGAT 1980
Qy 1981 CGCAAGCAATGGGCGTGCAGGCAAGTATTCCTCGAACTTTGGCAGACTGCGGATAAGA 2040
Db 1981 CGCAAGCAATGGGCGTGCAGGCAAGTATTCCTCGAACTTTGGCAGACTGCGGATAAGA 2040
Qy 2041 ATAAAAAGAAATTTGGTAAAGCTTCTGAAAGATGCAAAAGATAACGATTAATTTCCGATTA 2100
Db 2041 ATAAAAAGAAATTTGGTAAAGCTTCTGAAAGATGCAAAAGATAACGATTAATTTCCGATTA 2100
Qy 2101 AGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACCGCGTGAAGGATATA 2160
Db 2101 AGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACCGCGTGAAGGATATA 2160
Qy 2161 TCGGCTCTGTGGAGAGGCGAAGATCGCGGACTTTGGTGGTGAATCTCTGCTTTTGTG 2220
Db 2161 TCGGCTCTGTGGAGAGGCGAAGATCGCGGACTTTGGTGGTGAATCTCTGCTTTTGTG 2220
Qy 2221 GCGTAAACCAAAATCGTGATCAAGCGGCTATGCTCTTCTCTGAAATGGGCGAT 2280
Db 2221 GCGTAAACCAAAATCGTGATCAAGCGGCTATGCTCTTCTCTGAAATGGGCGAT 2280
Qy 2281 CTAAACCGCTCTGTGCCCACTCCCAACCGGTTTATTAACCGGAAATGTTTGGGCACTACG 2340
Db 2281 CTAAACCGCTCTGTGCCCACTCCCAACCGGTTTATTAACCGGAAATGTTTGGGCACTACG 2340
Qy 2341 GCAAGGCAAAATTTGACACCGATCACTTTTGTGTTTCAAAAGTCCGCTATGAAAAATGGCG 2400
Db 2341 GCAAGGCAAAATTTGACACCGATCACTTTTGTGTTTCAAAAGTCCGCTATGAAAAATGGCG 2400
Qy 2401 TGAAGAAAGCTGGGCTTAGAGCGCAAGTCTTACCGGTCAAAACCTGCCGTAAACATCA 2460
Db 2401 TGAAGAAAGCTGGGCTTAGAGCGCAAGTCTTACCGGTCAAAACCTGCCGTAAACATCA 2460

```
QY 2461 CCAAGAAAGACTTCAAGTTCAACGACAAACCGGCAAAAATCACCGTCGATCCGAAACCT 2520
Db 2461 CCAAGAAAGACTTCAAGTTCAACGACAAACCGGCAAAAATCACCGTCGATCCGAAACCT 2520
QY 2521 TCGAGGTCTTTGTAGATGGCAACTCTGCACCTCTAAACCCACCTCGCAAGTCCTCTAG 2580
Db 2521 TCGAGGTCTTTGTAGATGGCAACTCTGCACCTCTAAACCCACCTCGCAAGTCCTCTAG 2580
QY 2581 CCCAGCGCTACACTTTCTTAGGCAACATGCCCCCTTTGGGGGCAAGGTTATTTTAGGAA 2640
Db 2581 CCCAGCGCTACACTTTCTTAGGCAACATGCCCCCTTTGGGGGCAAGGTTATTTTAGGAA 2640
QY 2641 TCTTCATCAACGCACTCGCATCGGTCTTGCGTGTGCGATCGTGTCTTTAAACCAAC 2700
Db 2641 TCTTCATCAACGCACTCGCATCGGTCTTGCGTGTGCGATCGTGTCTTTAAACCAAC 2700
QY 2701 TTTTCATCTTTAAGCAATCGCATTTTAAATTAATTTAAATTTAATTAATTAATTAAT 2760
Db 2701 TTTTCATCTTTAAGCAATCGCATTTTAAATTAATTTAAATTTAATTTAATTAATTAAT 2760
QY 2761 TATGCCCTCTCATTTTAAAGGAGAAATATGCTAGTCTTTGGTATTCTATTAATTAATTAAT 2820
Db 2761 TATGCCCTCTCATTTTAAAGGAGAAATATGCTAGTCTTTGGTATTCTATTAATTAATTAAT 2820
QY 2821 TTGTTTGGTGTGGGCGCAAGGGTATTGAAACCCATCGCTCAAAAAGTAGAAGCCAC 2880
Db 2821 TTGTTTGGTGTGGGCGCAAGGGTATTGAAACCCATCGCTCAAAAAGTAGAAGCCAC 2880
QY 2881 AGG 2883
Db 2881 AGG 2883
```

RESULT 2

```
US-09-904-994B-13
; Sequence 13, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 2452
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(728)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (739)..(2445)
US-09-904-994B-13
```

Query Match 77.3%; Score 2228; DB 3; Length 2452;

Best Local Similarity 94.3%; Pred. No. 0;

Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

```
QY 159 AGGACTATTAATCCCATTTGCCCTTTAAATTTAAACAAAGGAGTAATAGGTGAACCTCACAC 218
Db 1 AGGACTATTAATCCCATTTGCCCTTTAAATTTAAACAAAGGAGTAATAGGTGAACCTCACAC 60
QY 219 CCAAGAGCAAGAAAGTTCTTGTTATATTATGCGGCGCAAGTGGCTAGAAAGCGCAAG 278
Db 61 CCAAGAGCAAGAAAGTTCTTGTTATATTATGCGGCGCAAGTGGCTAGAAAGCGCAAG 120
QY 279 CAGAGGGCTTAAAGCTCAACCAACCGCAAGCCATTGCTTTACATTAGTGCCCATATTATGG 338
Db 121 CAGAGGGCTTAAAGCTCAACCAACCGCAAGCCATTGCTTTACATTAGTGCCCATATTATGG 180
```

```
QY 339 ACGAAGCGCGCCCTGGGAAAAAACCCTTGCAGCTTATGGAAGAGTGCATGCACCTTTT 398
Db 181 ACGAGCGCGCTCGTGGCAAAAAACCCTTGCAGCTTATGGAAGAGTGCATGCACCTTTT 240
QY 399 TGNAAAAAGATGAAGTAATGCGCGGGTGGTAAATATGTTCCCGATCTAGGTGTAGAG 458
Db 241 TGNAAAAAGATGAAGTAATGCGCGGGTGGGAAATATGTTCCCTGATTTTGGCGGTGGAAG 300
QY 459 CCACCTTTCTGTAGTGTACGAAACTTGTAACTGTGAAATTTGGCCCATCGAACCATGAGC 518
Db 301 CCACCTTTCCCGATGGCACCAACTCGTAACTGTGAAATTTGGCCCATCGAACCATGAGC 360
QY 519 ACTTCAAGCGCGCAAGTGAATTTGGTTGCAATAAAGACATCGAGCTCAATGCAAGCA 578
Db 361 ACTTTAAGCGCGGTGAAGTGAATTTGGCTGTGATAAAGACATTTGAATCAACCGCAGTA 420
QY 579 AAGAAGTAACCGCAACTTGAAGTTACTAAATGAAGGGCTTAATCTTTGCATGTGGGTAGCC 638
Db 421 AAGAAGTTACCGAACTAGAAAGTTACTAAACGAAGGACCTAAATCTTTGCATGTGGGTAGCC 480
QY 639 ATTTCCACTTTCTTTGAAGCTTAAACAGGCCTAAATTTGATCTGTAATAAAGCCCTTATGCA 698
Db 481 ATTTCCACTTTCTTTGAAGCAAAAGCATTGAATTTGATCGGAAATAAAGCCCTTATGCA 540
QY 699 AACGCTTAGATATTTCCCTCTGCGCAACACGCTACGCAATTTGGGCGAGGCAAAACCCGCAAG 758
Db 541 AACGCTTAGATATTTCCCTCTGCGCAACACGCTACGCAATTTGGGCGAGGCAAAACCCGTAAG 600
QY 759 TGCAGTTGATTTCTTTGGTGGCAGTAAATAAGTGAATTTGCAATGACGGCTTGTGAATA 818
Db 601 TGCAGTTAATTTCTTTGGCGGTAGTAAATAAGTGAATTTGCAATGACGGCTTGTGAATA 660
QY 819 ACATCGCGGATGAACGCCATATAACATTAAGCGCTTGCACAGGCGAAATCTCACGGATTTA 878
Db 661 ATATTGCGGACGAACGCCATATAACATTAAGCGCTTGCACAGGCGAAATCTCACGGATTTA 720
QY 879 TCAAGTAAGGAGACTCCCATGAAAAATCAAAAAACAAGAAATATGTAATACTCTACGGACCC 938
Db 721 TCAAGTAAGGAGACTCCCATGAAAAATGMAAAACAAGAGTATGTAATACTCTACGGACCC 780
QY 939 ACCAAAGCGGATAAAGTGGCTTTAGAGATACCAGTATTTGGGCGAGAGTAGAACAATGAC 998
Db 781 ACCAAGCGGATAAAGTGGCTTTAGAGATACCAGTATTTGGGCGAGAGTAGAACAATGAC 840
QY 999 TATACCACTATGGCGAAGCACTTAAATTTGGCGCGGTAAACTATCCGTGAGGGTATG 1058
Db 841 TATACCACTATGGCGAAGCACTTAAATTTGGCGCGGTAAACTATCCGTGAGGGTATG 900
QY 1059 GGTGAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTTAGTCACTAACCGCATGATT 1118
Db 901 GGTGAGAGCAATAGCCCGAGATGAAACACCTTAGATTTAGTGATCACCACCGCATGATT 960
QY 1119 ATGCACTACACCGGATTTTCAAAAGCGCAATTTGGGATTTAAAAACGGCAAAATCCATGGC 1178
Db 961 ATTGACTACACCGGATTTTCAAAAGCGCAATTTGGCATTAAAAATGGCAAAATCCATGGC 1020
QY 1179 ATTGGCAAGCGAGGAAACAAAGGATGCAAGATGCGGTAGCCCTCATATGGTCTGGGT 1238
Db 1021 ATTGGCAAGCGAGGAAACAAAGGATGCGGTAGCCCTCATATGGTCTGGGT 1080
QY 1239 GTGGGCACAGAGCACTAGCAGGGGAGGTATGATTTATTTACCGTGGGGAAATCGATTCA 1298
Db 1081 GTGGGCACAGAGCACTAGCAGGGGAGGTATGATTTATTTACCGTGGGGGATCGATTCA 1140
QY 1299 CACACCACTTCTCTTCTCCACAAATTTCCCTACCGCTTAGCCAAATGGCGTTTACAACC 1358
Db 1141 CACACCACTTCTCTCTCCACAAATTTCCCTACCGCTTAGCCAAATGGCGTTTACAACA 1200
QY 1359 ATGTTTGGAGGGGACAGGTCTGTAGATGGCAAGATGCGACTACTATCACTCCGGGC 1418
Db 1201 ATGTTTGGCGGTGGCACAGGCCCGGTAGATGGCAAGATGCGACTACTATCACTCCGGGC 1260
QY 1419 AAATGGAACTTGCACCGCATGTTTGGCGGACGAGAGATTTCTTATGAATGTGGGCTTT 1478
```

Db 1261 AAATGGAACCTTGACCGCATGTTGCGCGCAGAGAGATATTCATGAATGTTGGGCTTT 1320
Qy 1479 TTGGGCAAGCAATAGCTTAGCAAAAAACAACCTTGTAGAAACAAGTAGAGGGGCGCG 1538
Db 1321 TTGGGCAAGCAATAGCTTAGTAAAAACAACCTTGTAGAACAAAGTAGAGGGGCGCG 1380
Qy 1539 ATTGGTTTAAATGTGATGAAGACTGGGGCAACAACCAAGTGTGATCGATCACTGTGTTG 1598
Db 1381 ATTGGTTTAAATGTGATGAAGACTGGGGCAACAACCAAGTGTGATCGATCACTGTGTTG 1440
Qy 1599 AGGTGCGAGATCAATACGATGTCAGTTTGTATCCACACCGATACAGTCAATGAGGCA 1658
Db 1441 AGGTGCGAGATCAATACGATGTCAGTTTGTATCCACACCGATACAGTCAATGAGGCA 1500
Qy 1659 GGTATTGTAGATGACACCTTAAATGCAATGAAGGGGCGCGCATTCATGCTTACCAATTT 1718
Db 1501 GGTATTGTAGATGACACCTTAAATGCAATGAAGGGGCGCGCATTCATGCTTACCAATTT 1560
Qy 1719 GAGGAGCGGTGGAGACACTCACCTGTATATCAACATGGCAGCGAGCTCAATATT 1778
Db 1561 GAGGAGCGGTGGAGACACTCACCTGTATATCAACATGGCAGCGAGCTCAATATT 1620
Qy 1779 CTACCTCTCCACCCACCTATTCCTATACCAATTAATACGTTGCGAGACACTTA 1838
Db 1621 CTACCTCTCCACCCACCTATTCCTATACCAATTAATACGTTGCGAGACACTTA 1680
Qy 1839 GACATGCTCATGATGCCACCACTAGACAAACCGCATCCGCGAGGATTTCAATTTTCT 1898
Db 1681 GACATGCTCATGATGCCACCACTAGATAAACGATCCGCGAGGATTTCAATTTTCT 1740
Qy 1899 CAAAGCGTATCCGCGCGCTCTATCGCGCTGAAGATGTCTCCATGATATGGGTG 1958
Db 1741 CAAAGCGTATCCGCGCGCTCTATCGCGCTGAAGATGTCTCCATGATATCGGTG 1800
Qy 1959 ATCGCATGACAGCTCGGATTCGCAAGCAATGGGGCGTGCAGCGAGTATCTCTGA 2018
Db 1801 ATCGCATGACAGTTCGATTCGCAAGCAATGGGGCGCTGGGGAAGTATTCCTAGA 1860
Qy 2019 ACTTGGCAGACTCGGATAGAAATAAAAAGAAATTTGGTAAGCTTCTGGAAGTGGCAAA 2078
Db 1861 ACTTGGCAGACTCGGATAGAAATAAAAAGAAATTTGGTAAGCTTCTGGAAGTGGTGA 1920
Qy 2079 GATAACGATATTTCCGCTATTAAGCGCTACATCTCAAATAACATATCAACCCCGCTTTG 2138
Db 1921 GATAATGACAACTTCGCAATCAAAAGCTATATCTCAAATAACATATTAATCCCGCTTTG 1980
Qy 2139 ACCACGCGTGAAGCGATATATCGGCTCTGTGGAAGAGGCAAGATCGCGACTTGGTG 2198
Db 1981 ACCATGCGCTGAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCGACTTGGTG 2040
Qy 2199 GTGTGGAATCCTGCTTTTGGCGTAAACCCCAAAATCGTGATCAAGGCGGTATCGTG 2258
Db 2041 GTGTGGAATCCTGCTTTTGGCGTAAACCCCAAAATCGTGATCAAGGCGGTATCGTG 2100
Qy 2259 GTCTTCTCTGAAATGGCGGATTTCTAAACGCTCTGTGCCACTCCGCCAACCGGTTTATTAC 2318
Db 2101 GTGTCTCTGAAATGGCGGATTTCTAATGCGTCTGTGCCACTCTCAGCGGTTTATTAC 2160
Qy 2319 CGGAAATGTTTGGGCAATCAAGGCAAGGGAATTTGACACCAAGATCACTTTTGTTC 2378
Db 2161 CGGAAATGTTTGGGCAATCAAGGCAAGGGAATTTGACACCAAGATCACTTTTGTTC 2220
Qy 2379 AAAGTCCGCTATGAAATGGGTGAAGAAAGCTGGGCTTAGAGCGCGAAGTGTCTCCCG 2438
Db 2221 AAAGTCCGCTATGAAATGGGTGAAGAAAGCTAGGTTTAGAGCGCGAAGTGTCTCCCG 2280
Qy 2439 GTCAAAACCTCGCGTAACTACCAAGAAAGCTTCAAGTTTCAACGCAAAACCGGCAAAA 2498
Db 2281 GTGAAACCTCGCGTAACTACCAAGAAAGCTTCAAGTTTCAACGCAAAACCTGCAAAA 2340
Qy 2499 ATCACCGTGCATCGGAAAACTTTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2558

Db 2341 ATCACCGTGCATCGGAAAACTTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2400
Qy 2559 CCCACCTCGCAAGTGCCTCTAGCCAGCGCTACACTTTTCTTAGGCAAAAT 2610
Db 2401 CCCACCTCTGAAGTGCCTCTAGCCCAACGCTACACTTTTCTTAGGCATAAT 2452

RESULT 3
US-09-904-994B-4
; Sequence 4, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904.994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2405
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(681)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (692)..(2398)
US-09-904-994B-4

Query Match 74.9%; Score 2158.6; DB 3; Length 2405;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

Qy 206 GTGAACTCTACACCAAGAGAGGCTTAAAGCTCAACCAACCCGAAAGCCATTGCTTACATTAGT 265
Db 1 GTGAACTCTACACCAAGAGAGGCTTAAAGCTCAACCAACCCGAAAGCCATTGCTTACATTAGT 60
Qy 266 AGAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCCGAAAGCCATTGCTTACATTAGT 325
Db 61 AGAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCCGAAAGCCATTGCTTACATTAGT 120
Qy 326 GCCCATATTATGACGAGCGCGCTGGAAAAAACCCTTGGCCAGCTTATGGAAGAG 385
Db 121 GCCCATATTATGACGAGCGCGCTGGAAAAAACCCTTGGCTGAACCTTATGGAAGAA 180
Qy 386 TGCATGACATTTTGAAGAAAGATGAAGTAATGCGCGGGTGGTAAATATGGTTCCCGAT 445
Db 181 TGTATGACATTTTGAAGAAAGATGAGTGTATGCGCGGTGGGGAATATGGTCCCTGAT 240
Qy 446 CTAGGTGTAGAACCCACTTTCTCTGATGTAACGAAACTTGTAACTGTGAAATTTGGCCCATC 505
Db 241 TTGGGCGTAGAACCCACTTTCCCGCATGGCACCAACTCGTAACTCGTAATTTGGCCCAT 300
Qy 506 GAAACAGATGACACTTCAAGCGGCGAAGTGAATTTGGTTGCGATAAAGACATCGAG 565
Db 301 GAACTGTAGAACACTTTAAAGCGGTGAAGTGAATTTGGCTGTGATAAAGACATTTAG 360
Qy 566 CTCAATGAGCAAGAAAGTAAACCGAATTTGAGGTACTAATGAAGGCGCTTAATCTTTG 625
Db 361 CTCAACGCGGTAAAGAAAGTTACCGAGCTTGAAGTTTACCAACGAAAGGACCTTAAATCTTTG 420
Qy 626 CATGTGGTAGCCATTTCCACTTCTTTGAAGCTAACAGGCACTAAAAATTCGATCGTAA 685
Db 421 CATGTGGTAGCCATTTCCACTTCTTTGAAGAACCAACAGGCAATTTGAATTTGATCGGAA 480
Qy 686 AAAGCCTATGGCAACCGCTAGATATTTCCCTCTGGCAACACGCTACGATTTGGGGCAGGA 745
Db 481 AAAGCCTATGGCAACCGCTAGATATTTCCCTCTGGCAACACGCTACGATTTGGGGCAGGA 540
Qy 746 CAAACCGGCAAGTGCAGTTGATTTCTTGGTGGCAGTAAATAAGTGAATTCGATGAAC 805

Db 541 CAAACCGGTAAAGTCAGTAAATCCCTCTTTGGCGGTAGTAAATAAGTGAATGGCATGAAC 600
Qy 806 GGGCTTTGTAATAACATCGCGATGAACGCCATAAATCAATAAAGCGCTTGACAAAGGCGAAA 865
Db 601 GGGCTTTGTAATAATTTGGGACGAAGCCATAAATCAACAAAGCACTAGACAAAGGCAAAA 660
Qy 866 TCTCAGCGAATTTCAAGTAAGAGACTCCCATGAAATGAAATAAACAAGATATGTATAA 925
Db 661 TCTCAGCGAATTTCAAGTAAGAGACTCCCATGAAATGAAATAAACAAGATATGTATAA 720
Qy 926 TACCTACGGACCCACCAAGGGGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCAGA 985
Db 721 CACCTACGGACCCACCAAGGGGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCAGA 780
Qy 986 AGTAGAACATGACTATACCACTCTATCGCGAAGAACTTAAATTTGGCGCGGTAAAACTAT 1045
Db 781 AGTAGAACATGACTATACCACTCTATCGCGAAGACTTAAATTTGGCGCGGTAAAACTAT 840
Qy 1046 CGGTAGGGTATGGGTGAGAGCAATAGCCCTGATGAAAAACACCTAGATTTAGTCATCAC 1105
Db 841 CGGTAGGGTATGGGTGAGAGCAATAGCCCTGATGAAAAACACCTAGATTTAGTCATCAC 900
Qy 1106 TAACCGATGATATGACTACACCGGATTTACAAAGCGGACATTTGGGATTAATAAACGG 1165
Db 901 CAAACCGATGATATGACTACACCGGATTTATAAAGCGGACATTTGGTATTAAAAAATGG 960
Qy 1166 CAAATTCATGGATTGGCAAGGACAGAAACAGGACATGCAAGTGGCGTAAAGCCCTCA 1225
Db 961 CAAATTCATGGATTGGCAAGGACAGGAGAAACAGACATGCAAGTGGCGTAAAGCCCTCA 1020
Qy 1226 TATGTCGTGGGTGGGCACAGAGCACTAGCAGGGGAAGGTATGATTTATTACCGCTGG 1285
Db 1021 TATGTCGTGGGTGGGCACAGAGCACTAGCAGGGGAAGGTATGATTTATTACCGCTGG 1080
Qy 1286 GGAATTCGATTACACACCCACTTCTTCTTCCACAAATTTCCCTACCGCTCTAGCCAA 1345
Db 1081 GGGGATTCGATTTCGACACCCACTTCTTCTTCCCAAAATTTCCCTACCGCTCTAGCCAA 1140
Qy 1346 TGGCGTTACACCATGTTTGGAGGGGACAGCTCTGTAGATGGCAGCAATGCGACTAC 1405
Db 1141 TGGTGTTACACCATGTTTGGAGGTGGCACAGTCCGGTAGATGGCAGCAATGCGACCAC 1200
Qy 1406 TATCACTCCGGGCAATGGAATTCGACCGCATGTTGCGCGCAGCAGAGAGATTTCTAT 1465
Db 1201 CATCACTCCGGCAATGGAATTCGACCGCATGTTGCGCGCAGCTGAAGAGTATTTCTAT 1260
Qy 1466 GAAATGTCGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATTTGTAGAACAGT 1525
Db 1261 GAAATGTCGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATTCGTAGAACAGT 1320
Qy 1526 AGAAGCGGCGCATTTGGTTTAAATTTGCATGAAGCTGGGGCACAACACCAAGTCCGAT 1585
Db 1321 AGAAGCGGCGCATTTGGTTTAAATTTGCATGAAGCTGGGGCACAACCAAGTCCGAT 1380
Qy 1586 CGATCACTCTTGAGCGTGCGAGATGAATPACGATGCAAGTTTGTATCCACACCGATAC 1645
Db 1381 CGATCACTCTTGAGCGTGAGAGATGAATPACGATGCAAGTTTGTATCCACACCGATAC 1440
Qy 1646 AGTCAATAGGCGAGGTTATGTAGATGACACCTTAAATGCAATGAACGGCGCGCCATCCA 1705
Db 1441 GGTCAATAGGCGAGGTTATGTAGATGACACCTTAAATGCGATGAACGGCGCGCCATCCA 1500
Qy 1706 TGCCTACCACTTAGGAGCGGGTGGAGGACACTCACCTGATGTTTATCACCATGTCAGG 1765
Db 1501 TGCCTACCACTTAGGAGCGGGTGGAGGACACTCACCTGATGTTTATCACCATGTCAGG 1560
Qy 1766 CGAGCTCAATATTTTACCTCTCTCCACACCCCTATTTCCCTATPACCATTAATACGCT 1825
Db 1561 CGAGCTCAATATTTTACCTCTCTCCACACCCCTATTTCCCTATPACCATTAATACGCT 1620
Qy 1826 TGCAGAACACTTAGACATCTCATGACATGCAACCACTAGACCAACCGATCCGCGAGGA 1885
Db 1621 TGCAGAACACTTAGACATCTCATGACCTGCCACACCTAGACCAACCGATCCGCGAGGA 1680

Qy 1886 TTTCACAAATTTTCTCAAAGCCGTATCCGCCCGGCTCTATCCCGCTGAAGATGTGCTCCA 1945
Db 1681 TCTCCAGTTTTCCAAAGCCGTATCCGCCCGGCTCTATTTGCCGTGAAGATGTGCTCCA 1740
Qy 1946 TGNATGGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGTCAGGCGA 2005
Db 1741 TGNATTTGGCGTATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGTCGCGGA 1800
Qy 2006 AGTGATTTCTCGAACTTTGGCAGACTGCGGATAAGATAAAAAAGAAATTTGGTAGCTTCC 2065
Db 1801 AGTGATTTCTAGAACTTTGGCAAACTGCAGACAGAAATAAAAAGAAATTTGGTAGCTTCC 1860
Qy 2066 TGAAGATGGCAAGATAACGATAATTTCCGATTTAAGCGGTATACATCTTCAAATACATAT 2125
Db 1861 TGAAGATGGTCAGATAATGACAACTTCGCGATCAAACGCTATATCTTCAAATACACCAT 1920
Qy 2126 CAACCCCGCTTTGACCCACCGCGTACGAGTATATCGGCTCTGTGGAAGGGCAAGAT 2185
Db 1921 TAAATCCCGCTTTGACCCATCGCGTACGAGTATATCGGCTCTGTGGAAGGGCAAGAT 1980
Qy 2186 CGCCGACTTGGTGTGGGAATCTGCTTTTGGCGTAAACCCTAAATTCGTGATCAA 2245
Db 1981 CGCCGACTTGGTGTGGGAATCTGCTTTCTTTGGTGTAAACCCTAAATTCGTGATCAA 2040
Qy 2246 AGCGGTATGGTGTGCTTCTCTGAAATGGCGGATTTCTAACCGCTCTGTGCCCACTCCCCA 2305
Db 2041 AGCGGTATGGTGTGCTTCTCTGAAATGGCGGATTTCTAACCGCTCTGTGCCCACTCCA 2100
Qy 2306 ACCGTTTATTTACCGGCAATTTTGGGCATTCACGGCAAGCGGCAAAATTTGACACCAAGAT 2365
Db 2101 GCGGTTTATTTACCGGCAATTTTGGGCATTCACGGCAAGCGGCAAAATTTGACACCAAGAT 2160
Qy 2366 CACTTTTGTTCCAAAGTCGCTATGAAATGCGTGAAGAAAGCTGGGCTTAGAGCG 2425
Db 2161 CACTTTTGTTCCAAAGTCGCTATGAAATGCGGTAAGAAAGAACTAGGCTTAGAGCG 2220
Qy 2426 CCAAGTTCTACCGGTCAAAACTGCGGTAACTACCAAGAAAGACTTTCAAGTTCAACGA 2485
Db 2221 CCAAGTTCTACCGGTCAAAACTGCGGTAACTACCAAGAAAGACTTTCAAGTTCAACGA 2280
Qy 2486 CAAAACGGCAAAATCACCGTCGATCGAAAACTTCGAGGTCTTTGTAGATGGCAAACT 2545
Db 2281 CAAAGCGGCGCATATCACTGTCGATCTTAAACCTTCGAGGTCTTTGTAGATGGCAAACT 2340
Qy 2546 CTGCACCTTAACCCCACTCGCAAGTCCCTTAGCCCGGCTACACTTTTCTTAGGC 2605
Db 2341 CTGCACCTTAACCCCGCTCTGAAGTGCCTTAGCCCAACGCTACACTTTTCTTAGGC 2400
Qy 2606 ACAAT 2610
Db 2401 ACAAT 2405

RESULT 4
US-09-904-994B-10
; Sequence 10, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 2407
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(682)

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (693) .. (2399)
US-09-904-994B-10

Query Match      74.9%; Score 2158.6; DB 3; Length 2407;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy 206 GTGAAGTCTCACACCCCAAGAGCAGCAAGAAAAGTCTCTGTATATATATGCGGCGAAGTGGCT 265
Db 2 GTGAAGTCTCACACCCCAAGAGCAGCAAGAAAAGTCTCTGTATATATATGCGGCGAAGTGGCT 61

Qy 266 AGAAGCGGCAAGCAGAGGCTTAAAGCTCAACCAACCCGAGCCATTTGCTTACATAGT 325
Db 62 AGAAGCGCAAGCGGAGGCTTAAAGCTCAACCAACCCGAGCCATTTGCTTACATAGT 121

Qy 326 GCCCATATATGACCAAGCGCGCTGGAAAAAAGCAACCGTCCCAAGCTTATGGAAGAG 395
Db 122 GCCCATATATGACCAAGCGCGCTGGAAAAAAGCAACCGTCCCAAGCTTATGGAAGAG 181

Qy 386 TGCATGCACCTTTTGAAGAAAGATGAAGTAATGCCCGGGTGGGTAAATATGTTTCCCGAT 445
Db 182 TGTATGCACCTTTTGAAGAAAGACGAGGTGATGCCCGGTGGGGAAATATGTTCCCTGAT 241

Qy 446 CTAGGTGTAGAGCCACCTTTCTGTATGTGTAACGAAAATTGTGTAATCTGTGAATTTGGCCCATC 505
Db 242 TTAGGGGTGGAAGTACTTTTCCGATGGCACCACCAACTCTGTAACCCGTGAATTTGGCCCATC 301

Qy 506 GAACCAAGTACGACCTTCAAGCGGCGGAGTGAATTTGGTTGCGATGAAGACATCGAG 565
Db 302 GAACCAAGTGAACACTTTCAAGCGGCGGAGTGAATTTGGTTGCGATGAAGACATTTGAA 361

Qy 566 CTCATATGCAAGCAAGAGTAACCGAATTTGAGGTTACTAATGAAGGGCTTAATATCCTTG 625
Db 362 CTCATATGCAAGTAAGNAGTTACCGAATAGAGTTACCAACGAGACCTTAATCCTTG 421

Qy 626 CATGTGGGTAGCCATTTCCACTTTTGAAGCTTAAAGGCACTAAATTCGATCGTGAA 685
Db 422 CATGTGGGTAGCCATTTCCACTTTTGAAGCCCAACGAGCAATTTGAAATTCGATCGGAA 481

Qy 686 AAGCCATATGGCAAGCCCTAGATATTCCTCTGGCAACAGCTACGCATTTGGGGCAGGA 745
Db 482 AAGCCATATGGCAAGCCCTAGATATTCCTCTGGCAACAGCTACGCATTTGGGGCAGGA 541

Qy 746 CAACCCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAAAGTGAATTTGGCATGAAC 805
Db 542 CAACCCGCAAGTGCAGTTGATTCCTCTTGGGCGCAGTAAAAAGTGAATTTGGCATGAAC 601

Qy 806 GGGCTTTGTGAATTAACATCGCGGATGAACGCGCATAAACATAAAGCGCTTGAACAGGCGAAA 865
Db 602 GGGCTTTGTGAATTAATTTGCAGATGAACGCGCATAAACATAAAGCGTTTGAAGAAAAGCAAAA 661

Qy 866 TCTCAGGATTTATCAAGTAAGAGACTCCCATGAAATGAAAAACAAAGATATGTAAG 925
Db 662 TCTCAGGATTTATCAATTAAGAGACTCCCATGAAATGAAAAACAAAGATATGTAAG 721

Qy 926 TACCTACGACCCCAACCAAGCGCAAGTAAAGTGCCTTAGGAGATACCCGATCTTTGGGCAGA 985
Db 722 TACCTACGACCTTACCAAGCGCAAGTAAAGTGCCTTAGGAGATACCCGATCTTTGGGCAGA 781

Qy 986 AGTAGAACATGATATACCACTATATGCGGAAGAACTTTAAATTTGGCGCGGGTAAAACTAT 1045
Db 782 AGTAGAACATGATATACCACTATATGCGGAAGAGCTCAATTTGGCGCGGGTAAAACTAT 841

Qy 1046 CCGTAGGGTATGGGTACAGAGCAATAGCCCTGTATGAAAAACCCCTAGATTTAGTCATC 1105
Db 842 CCGTAGGGCATGGGTACAGAGCAATAGTCCAGATGAAAAACCCCTAGATTTAGTCATC 901

Qy 1106 TAAGCGGATGATTCGACTACACCGGATTTACAAAGCCGACATTTGGGATTAAGAACGG 1165
Db 902 CAACCGGATGATTTAGTACACCGGATTTTCAAAAGCCGACATTTGGCATTAAGAACGG 961
```

```
Qy 1166 CAAAATCCATGGCATTTGGCAAGCGCAAGAAACAAGGACATGCAGATCGCGTAAGCCCTCA 1225
Db 962 CAAAATCCATGGCATTTGGCAAGCGCAAGAAACAAGGACATGCAGATGCAGTAAGCCCTCA 1021

Qy 1226 TATGSGTCGTGGGTGTGGGCACAGAACGACTAGCAGGGGAAGGTATGATTTATACCGCTGG 1285
Db 1022 TATGSGTCGTGGGTGTGGGCACAGAACGACTAGCAGGGGAAGGTATGATTTATACCGCTGG 1081

Qy 1286 GGAATTCGATTTCAACACCCACTTCTTTCTCCACAACAATTTCCCTACCGCTCTAGCCAA 1345
Db 1082 GGGGATCGATTTCAACACCCACTTCTCTCTCCACAACAATTTCCCTACCGCTCTAGCCAA 1141

Qy 1346 TGGCGTTTACAAACCATTTTGGAGCGGCACAGGTCCTGTAGTGGCAAGATGGGACTAC 1405
Db 1142 TGGCGTTTACAAACCATTTTGGCGGTGGCAAGGTCCTGTAGTGGCAAGATGGGACTAC 1201

Qy 1406 TATCACTCCCGGGCAAAATGGAACTTTGCACCCGATGTTGCGGGCAGCAGAGAAGTATTTCTAT 1465
Db 1202 CATCACTCCCGGGCAAAATGGAACTTTGCACCCGATGTTGCGGGCAGCTGAAGAGTATTTCTAT 1261

Qy 1466 GAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAACAACCTTGTAGAACAAAT 1525
Db 1262 GAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTCAGTAAAAACAACCTTGTAGAACAAAT 1321

Qy 1526 AGAAGCGGCGCGATTTGGTTTAAATTTGCATGAGACTGGGGCACACACCAAGTGGCAT 1585
Db 1322 AGAAGCGGCGCGATTTGGTTTAAATTTGCATGAGACTGGGGCACACACCTCAAGTGGCAAT 1381

Qy 1586 CGATCACTGCTTGAAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC 1645
Db 1382 CGATCACTGCTTGAAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC 1441

Qy 1646 AGTCAATAGGCGAGGTTATGTAGATGACACCTTAAATGCAATGAACGCGGCGGCCATCCA 1705
Db 1442 GGTCAATAGGCGAGGTTATGTAGATGACACCTTGAATGCGATGAACGCGGCGGCCATCCA 1501

Qy 1706 TGCTTACCACTTTCAGGAGGGGTGGAGCACACTCACTGATGTTATCCACATGGCGAGG 1765
Db 1502 TGCTTACCACTTTCAGGAGGGGTGGAGCACACTCACTGATGTTATCAACATGGCGAGG 1561

Qy 1766 CGAGCTCAATATTTCTACCTCTCCACACACCCCACTATTTCCCTATACCATTAATACGGT 1825
Db 1562 CGAGCTCAATATTTCTACCTCTCCACACACCCCACTATTTCCCTATACCATTAATACGGT 1621

Qy 1826 TGCAGAACACTTTAGACATGCTCATGACATGCCACCTAGATAAAACGATCCGCGAGGA 1885
Db 1622 TGCAGAACACTTTAGACATGCTCATGACATGCCACCTAGATAAAACGATCCGCGAGGA 1681

Qy 1886 TTTACAAATTTTCTCAAAAGCCGATTCGCGCCGGCTCTATCGCGGCTGAAGATGTGCTCA 1945
Db 1682 TTTACAAATTTTCTCAAAAGCCGATTCGCGCCGGCTCTATCGCGGCTGAAGATGTGCTCA 1741

Qy 1946 TGATATCGGTGTGATTCGCGATGACAAGCTCGGATTTCCGAAGCAATGGCGGCTGCAGGCGA 2005
Db 1742 TGATATTTGGCGTGTATTCGCGATGACAAGCTCGGATTTCCGAAGCAATGGCGGCTGCAGG 1801

Qy 2006 AGTGATTTCTCGAACTTTGGCAGACTGGCGGATAAGATAAAGAAATTTGGTAAGCTTCC 2065
Db 1802 AGTGATTTCTCGAACTTTGGCAGACTGGCGGATAAGATAAAGAAATTTGGTAAGCTTCC 1861

Qy 2066 TGAAGATGGCAAGATAACGATAATTTTCGCGATTTAAGCGCTACATCTCCAAATACACTAT 2125
Db 1862 TGAAGATGTGCAGATAACGAACTTTCCGATTAACAAACGCTTACATCTCCAAATACACTAT 1921

Qy 2126 CAACCCCGCTTTGACCCACCGCGTGAAGGATATCGGCTCTGGAAGAGGGCAAGAT 2185
Db 1922 TAACCCCGCTTTAACCCTATGGGGTAAGCGAGTATTCGGCTCTGTGGAAGAGGGCAAAAT 1981

Qy 2186 CGCGCATTTGCTGTGTGGAATCTCGCTTTTGGGCGTAAAAACCAAAATCGTGATCAA 2245
Db 1982 CGCTGATTTGCTGTGTGGAATCTCGCTTTTGGTGTGAAACCTAAGATTTGTGATCAA 2041

Qy 2246 AGCGCGTATGCTGTCTTCTCTGAAATGGGCGATTTCTAAACGCGCTGTGTGCCCACTCCCA 2305
```

```
Db 2042 AGCGGTTATGGTCTTCTGAAATGGCGACTCAACGCGTCCGTGCTACACCTCA 2101
Qy 2306 ACCGGTTTATTACCGGAAATGTTTGGGCATCACGGCAAGCGCAAAATTTTACACACGAGCAT 2365
Db 2102 GCGGTTTATTACCGGAAATGTTTGGGCATCACGGCAAGCGCAAAATTTTACACACGAGCAT 2161
Qy 2366 CACTTTTGTTCAAAGTCGCTTATGAAATATGCGGTGAAAGAAAAGCTTGGGCTTAGAGCG 2425
Db 2162 CACTTTTGTTCAAAGTCGCTTATGAAATATGCGGTGAAAGAAAAGCTTAGAGCG 2221
Qy 2426 CCAAGTTCTACCGGTCAAAAGCTCCGTAAACATCACCAAGAAAGACTTCAAGTTTCAACGA 2485
Db 2222 CAAAGTGTCTACCGGTCAAAAGCTCCGTCAACATCACTAAGAAAGACTTCAAAATTTCAACAA 2281
Qy 2486 CAAAACGGCAAAATCACCGTCGATCCGAAACCTTTCGAGGTCTTTGTAGATGCGAAACT 2545
Db 2282 CAAGCGGCGCATATCACTGTGTCATCTTAACCTTCGAGGTCTTTGTAGATGCGAAACT 2341
Qy 2546 CTGCACCTCTAAACCCACCTCGCAAGTGCCTCTAGCCAGCGCTACACATTTTCTTAGGC 2605
Db 2342 CTGCACCTCTAAACCCGCTCTGAAGTGCCTCTAGCCAGCGCTACACATTTTCTTAGGC 2401
Qy 2606 ACAATG 2611
Db 2402 NCAATG 2407

RESULT 5
US-09-904-994B-7
; Sequence 7, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2183
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(683)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (694)..(2181)
US-09-904-994B-7

Query Match 67.6%; Score 1948; DB 3; Length 2183;
Best Local Similarity 93.3%; Pred. No. 0;
Matches 2035; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 206 GTGAACCTCACACCCAAAGACGACGAGTCTTAAAGCTCAACCAACCCGAAAGCCATTGCTTACATTAGT 325
Db 3 GTGAACCTCACACCCAAAGACGACGAGTCTTAAAGCTTCTGTTATATATATGCGGCGGAGTGGCT 265
Qy 266 AGAAGCGCAAGACGAGGCTTAAAGCTCAACCAACCCGAAAGCCATTGCTTACATTAGT 325
Db 63 AGAAGCGCAAGACGAGGCTTAAAGCTCAATCAACCCGAAAGCCATTGCTTACATTAGT 122
Qy 326 GCCCATATTATGGAACGAGCGCGGTGGAAGAAAACCGTTGCCAGCTTATGGAAGAG 385
Db 123 GCCCATATTATGGAACGAGCGCGGTGGAAGAAAACCGTTGCTGAACTTATGGAAGAA 182
Qy 386 TGCATGCACTTTTGAAGAAAGATGAAGTAATGCCCCGGGGTGGTAAATATGTTTCCCGAT 445
Db 183 TGTATGCACTTTTGAAGAAAGATGAAGTATGCCCCGGGTGATGCCCCGGTATGTTCCCTGAT 242
Qy 446 CTAGGTGTAGAACCCACCTTCTCTGATGTACGAACTTGTAACTGTGAATTGGCCCCATC 505
```

```
Db 243 TTGGCGGTAGAAGCCACTTTCCCGATGGCACCAAACTCGTAAACGTTGAATTTGGCCCAATT 302
Qy 506 GAACCCAGATGAGCACTTTCAAAGCGGCGAAGTGAATTTTGGTTGCGATAAAGACATCGAG 565
Db 303 GAACCTGATGAACACTTTAAAGCGGTGAAGTGAATTTTGGCTGTATTAAGACATTTGAG 362
Qy 566 CTCAANTGACGCAAGAAAGTAAACGAACTTGAGGTTACTAATGAAGGCGCTAAATTCCTTG 625
Db 363 CTCAACGTGGGTAAAGGAAGTTACCGAGCTTGAAGTTTACCAACGAAGGACCTAAATTCCTTG 422
Qy 626 CATGTGGGTAGCCATTTTCCACTTTTGAAGCTTAACGAAGCACTTAAATTTGCGATCGTGAA 585
Db 423 CATGTGGGTAGCCATTTTCCACTTTTGAACCAACCAAGGCATTTGAATTTGATCGGGAA 482
Qy 686 AAAGCCCTATGGCAAAACCCCTAGATATTCCCTCTGGCAACACGCTACGCAATTTGGGGCAGGA 745
Db 483 AAAGCCCTATGGCAAAACCCCTAGATATTCCCTCTGGCAACACGCTACGCAATTTGGGGCAGGA 542
Qy 746 CAAACCCGCAAAAGTGCAGTTGATTTCTCTTTGGTGGCAGTAAAAAAGTGATTTGGCATGAAC 805
Db 543 CAAACCCGTAAGTGCAGTTAAATCCCTCTTTGGCGGTAGTAAAAAAGTGATTTGGCATGAAC 602
Qy 806 GGGCTTGTGAATTAACATCGCGGATGAAGCCATTAACATAAAGCGCTTGACAAGCGCGAAA 865
Db 603 GGGCTTGTGAATTAATTTGGCGAGCAACGCGCATTAACACAAGACACTAGACAAGGCAAAA 662
Qy 866 TCTCAGCGATTTTATCAAGTAAGGAGACTCCCATGAAATGAAGAAATGAAGAAATATGTAAA 925
Db 663 TCTCAGCGATTTTCAATCAAGTAAGGAGACTCCCATGAAATGAAGAAATGAAGAAATATGTAAA 722
Qy 926 TACTACGGACCCCAACCAAGCGCATAAAGTCGCTTAGGAGATACCGATCTTTGGGCGAGA 985
Db 723 CACTACGGACCCCAACCAAGCGCATAAAGTCGCTTAGGAGATACCGATCTTTGGGCGAGA 782
Qy 986 AGTAGAACATGACTATACCACTATGCGGAAGAACTTAAATTTTGGCGGGGTAAAACTAT 1045
Db 783 AGTAGAACATGACTATACCACTTATGCGGAAGAGCTCAAAATTTTGGCGGGGTAAAACTAT 842
Qy 1046 CCGTAGGGTATGGGTGAGAGCAATAGCCCTGTGAAAAACCCCTAGATTTAGTTCATCAC 1105
Db 843 CCGTAGGGTATGGGTGAGAGCAATAGCCCTGTGAAAAACCCCTAGATTTAGTTCATCAC 902
Qy 1106 TAACGCGATGATTATCGACTACACCGGATTTTCAAAAGCGACATTTGGGATTTAAAAACGG 1165
Db 903 CAACGCGATGATTATCGACTACACCGGATTTTAAAGCGACATTTGGTATTTAAAAATGG 962
Qy 1166 CAAAATCCATGGCAATTGGCAAGCGGAGAA CAAGGCATCAAGATCGCGTAAGCCCTCA 1225
Db 963 CAAAATCCATGGTATTGGCAAGCGGAGAA CAAGGCATCAAGATCGCGTAAGCCCTCA 1022
Qy 1226 TATGCTGTGGGTGTGGGCA CAGAGCACTAGCAGGGAAGGTATGATTTATTCGCTGG 1285
Db 1023 TATGCTGTGGGTGTGGGCA CAGAGCACTAGCAGGGAAGGTATGATTTATTCGCTGG 1082
Qy 1286 GGAATCGATTTCACACACCCACTTCCCTTCTCCACAACTTCCCTACCGCTCTAGCCAA 1345
Db 1083 GGGATCGATTTCGACACCCACTTCCCTTCTCCCAAACTTCCCTACCGCTCTAGCCAA 1142
Qy 1346 TGGCGTTACAAACCATTTTGGAGCGGCA CAGGTCCTGTAGTGGCA CGAATGCGACTAC 1405
Db 1143 TGGTGTACAAACCATTTTGGAGGTGGCA CAGGTCGCTAGATGGCA CGAATGCGACCAC 1202
Qy 1406 TATCACTCCGGGCAATTTGAACTTTGCA CCGCATGTTGCGGCGCAGCAGAAAGTATTTCTAT 1465
Db 1203 CATCACTCCGGGCAATTTGAACTTTGCA CCGCATGTTGCGGCGCAGCTGAAGAGTATTTCTAT 1262
Qy 1466 GATGTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAACCACTTGTAGAACAGT 1525
Db 1263 GAATGTAGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAACCACTTGTAGAACAGT 1322
Qy 1526 AGAAGCGGCGCGATTTGTTTAAATTTGAAGTGAAGCTGGGCGCAACACCAAGTGGCAT 1585
```



```
Db 1323 AGAAGCGGGCGGCGATTGGCTTTAAATTGCGTGAAGACTGGGGGCACAAACACCAAGTGGCAT 1382
QY CGATCACTGCTTGAGCGGTGGCAGATGAATACAGATGTGCAAGTTTGTATCCACACCGCATAC 1645
Db |||||
Db 1383 CGATCACTGCTTGAGCGGTGGCAGATGAATACAGATGTGCAAGTTTGTATCCACACCGCATAC 1442
QY 1646 AGTCAATAGGCGAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGGGCGCGCATCCA 1705
Db |||||
Db 1443 GGTCAATAGGCGAGGTTATGTGATGACACCCCTAAATGCAATGAACGGGGCGCGCATCCA 1502
QY 1706 TGCCTACCACTTAGGCGAGGTTGGAGGACACTCACCTGATGTTATACCATGCGAGG 1765
Db |||||
Db 1503 TGCCTACCACTTAGGCGAGGTTGGAGGACACTCACCTGATGTTATACCATGCGAGG 1562
QY 1766 CGAGCTCAATATTCTACCTCTCCACACCCCTCACTATTCCTTATACCAATTAATACGGT 1825
Db |||||
Db 1563 CGAGCTCAATATTCTACCTCTCCACACCCCTCACTATTCCTTATACCAATTAATACGGT 1622
QY 1826 TGCAGAACCTTAGACATGCTCATGACATGCCACCTTAGACAAACGGCATCCCGGAGGA 1885
Db |||||
Db 1623 TGCAGAACCTTAGACATGCTCATGACATGCCACCTTAGACAAACGGCATCCCGGAGGA 1682
QY 1886 TTTTACAAATTTCTCAAGCGGTATCCGCCCGGCTCTATCGCGCTGAAGATGTGCTCCA 1945
Db |||||
Db 1683 TTTTACAAATTTCTCAAGCGGTATCCGCCCGGCTCTATTCGCCGCTGAGGATGTGCTCCA 1742
QY 1946 TGATATGGGTGTGATCGCATGACAAAGCTCGAATTCGCAAGCAATGGGGCGGTGCGAGCGA 2005
Db |||||
Db 1743 TGATATGGGTGTGATCGCATGACAAAGCTCGAATTCGCAAGCAATGGGGCGGTGCGAGG 1802
QY 2006 AGTGATTCCTCGAACTTGGCAGACTCGGATAGAAATPAAAAAGAAATTTGGTAAAGTTCC 2065
Db |||||
Db 1803 AGTGATTCCTCGAACTTGGCAGACTCGGATAGAAATPAAAAAGAAATTTGGTAAAGTTCC 1862
QY 2066 TGAAGATGGCAAGATGAATTAATTCGCGATTAAGCGCTACATCTCCAAATACACTAT 2125
Db |||||
Db 1863 TGAAGATGGCGAGATGAACAACTTCGCGATCAAAACGCTATATCTCCAAATACACTAT 1922
QY 2126 CAACCCCGCTTGACCCACCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGAT 2185
Db |||||
Db 1923 TAATCCGCTTTGACCATGGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGAT 1982
QY 2186 CGCGCACTTGGTGTGGAATCTGCTTTTGGCGTAAAAACCCAAATCGTGATCAA 2245
Db |||||
Db 1983 CGCGCACTTGGTGTGGAATCTGCTTTTGGCGTGAACCTAAGATTGTGATTA 2042
QY 2246 AGCGGTATGGTGTCTTCTGAAATGGGCGATTTAAACGCTCTGTGCGCACTCCCA 2305
Db |||||
Db 2043 AGGTGGCATGGTGTCTTCTCGAAATGGGCGATTTCTAAACGCTCCGTGCGCCACGCTCA 2102
QY 2306 ACCGGTTTATTACCGCAATGTTTGGGCATCAGGCAAGCGAAATTTGACACCGCAT 2365
Db |||||
Db 2103 CCGGTTTATTACCGCAATGTTTGGGCACCGCAAGCGGAAATTTGACACCGCAT 2162
QY 2366 CACTTTTGTGTTTCCAAAGTCC 2385
Db |||||
Db 2163 CACTTTTGTGCTCTCAAGCG 2182
```

RESULT 6

US-10-639-273-1

; Sequence 1, Application US/10639273

; Publication No. US20040142343A1

; GENERAL INFORMATION:

; APPLICANT: Chang, Yung-Fu

; APPLICANT: Simpson, Kenneth W

; APPLICANT: Zhu, Jiaguan

; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES

; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS

; FILE REFERENCE: 19603/3881

; CURRENT APPLICATION NUMBER: US/10/639, 273

; CURRENT FILING DATE: 2003-08-12

; PRIOR APPLICATION NUMBER: 60/404,337

; PRIOR FILING DATE: 2002-08-16

; NUMBER OF SEQ ID NOS: 41

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 8407

; TYPE: DNA

; ORGANISM: Helicobacter bizzozeronii

US-10-639-273-1

```
Query Match 41.1%; Score 1184.4; DB 7; Length 8407;
Best Local Similarity 68.5%; Pred. No. 7.3e-260;
Matches 1681; Conservative 0; Mismatches 741; Indels 32; Gaps 2;

QY 183 AAATTAAACAAGAGTAATAGTGAAATCAACCCAAAGAGCAAGAAAGTCTTTGT 242
Db |||||
Db 2319 AAATTGGTAGAGGAGTTTAGGATGAAATTAACCCCTAAAGAGCTGACAAAGTCTCATGT 2378
QY 243 TATATTATGCGGCGAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAAC 302
Db |||||
Db 2379 TGCAATTATGCGGCGAATTTGGCTAAAAAAACGCAAGCAAAATGGCGTTAAAGTAAATTATA 2438
QY 303 CCGAAGCCATTGCTTACATTAGTGCCTCATATTATGGACGAAAGCGCGCTGGAAAAA 362
Db |||||
Db 2439 CTGAGGAGTAGCCCTCATCAGTGCCCATGTGATGGAGAAAGCCCGTGCGAGTAA 2498
QY 363 CCGTTGCCAGCTTATGGAAGAGTGCATGCATCTTTTGA 422
Db |||||
Db 2499 GTGTGGCGGATTTGATGCAAGAGGCGAGGACACTTCTTAAAGCTGATGATGTCTATGCCCG 2558
QY 423 GGGTGGTAAATATGGTTCCTGATCTAGTGTAGNAGCCACTTTCCTGATGGTACGAAC 482
Db |||||
Db 2559 GTGTAGCCCATATGATCCAGAAAGTGGGATTTGAAGCTAACTTCCCTGATGGGACAAAC 2618
QY 483 TTGTAACTGTGAATTTGGCCCATCGAACAGATGAGCACTTCAAAGCGGGCGAAGTAAAT 542
Db |||||
Db 2619 TGGTAACTTCCATACCCCGTTGAAGATGGTGGGCATAAATGGCTCCGGTGAAGTGA 2678
QY 543 TTGGTTCCGATAAAGACATCGAGCTCAATCGAGCAAGAGTAACACCGAACTTTGAGGTGA 602
Db |||||
Db 2679 TTTTGA 2738
QY 603 CTAATGAAGGCGCTAAATCTCTTGATGTGGTAGCCATTTCCACTTCTTTGAAGTAAACA 662
Db |||||
Db 2739 ATAA 2798
QY 663 AGGCACCTAAATTCGATCGTGA 722
Db |||||
Db 2799 AGCTTTTGG 2858
QY 723 ACAGCTACGCAATTTGGGGCAGGACAAACCCCGCAAGTGCAGTTGATTCCTCTTGGTGGCA 782
Db |||||
Db 2859 CCGCTGTGCGCTTTGAACCCCGTGA 2918
QY 783 GTAAAAAGTGAATGGCATGAACCGGGCTTGTGAATAACATCGGGATGAACGCCATAAAC 842
Db |||||
Db 2919 ACCA 2978
QY 843 ATAAAGCGCTTGA 877
Db |||||
Db 2979 AACTTGTCTCAAACGGCGCAAGAACATGGCTTTTGGTGTGTAATTTGGGTTCCGATA 3038
QY 878 ----ATCAAGTAAGGAGACTCCCATGAAAAATGAAAAA---ACAAGATATGTAAATACCT 930
Db |||||
Db 3039 AAAATTAAGGAAGGACATCCGATGA 3098
QY 931 ACGGACCCACCAAGGCGATAAAGTGGCTTTAGAGATACCGATCTTTGGGCGAGAGTAG 990
Db |||||
Db 3099 ATGACCCACATACGGCGATTAAGTGA 3158
QY 991 AACATGACTATACCACTATCGGAGNAACCTTAAATTTGGCGGGTAAACTATCCGTG 1050
Db |||||
Db 3159 AACATGACTGACCACTTATGCGGAAGAAATTAAGTTTGGTGGCGGTAAACCATTCGCG 3218
```


Db 187 AGCCAGAACTCGATCTTGTGCTCACTAAAGCCCTGATCGTGATTAACCCGGCATTTAT 246
Qy 1140 AAAGCCGACATTTGGGATTAATAACGCGCAAAATCCATGGCATTTGGCAAGGAGGAAACAAG 1199
Db 247 AAAGCCGATATTGGCATTAATAATGCGAAATCCATGGCATTTGGCAAGGAGGAAACAAG 306
Qy 1200 GACATGCAAGATGGGCTAAGCCCTCATATAGTGGTGGTGGTGGGCAAGAGCACTAGCA 1259
Db 307 GACATGCAAGATGGGCTTTCACAAATCTTTGGGTGGGCGCTGCTACTAGGGCTTTGGCC 366
Qy 1260 GGGGAGGTATGATTAATACCGCTGGGGAATCGATTCAACACCCACCTTCTTCTTCCA 1319
Db 367 GCTGAAGGGGTGATTTGACAGCTGGTGGGATTAACACCCACATCCACTTTATTTCTCC 426
Qy 1320 CAACAAATTCCTACCGCTCTAGCCAAATGGCGTTTACAAACCATGTTTGGAGCGGCACAGGT 1379
Db 427 CAACAAATCCCAACAGATTTGCGAGCGGATCACAACCATGATTTGGTGGGGAACAGGT 486
Qy 1380 CTTGTAGATGGCAAGATGGCACTATATCACTCGGGGCAAAATGGAACTTGGACCGCATG 1439
Db 487 CCAGCTGATGGGACTAAGCGCATTTCCAGCGGATCACAACCATGATTTGGTGGGGAACAGGT 546
Qy 1440 TTGGCGCAGCAGAGATTTCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCT 1499
Db 547 CTCGGTCCCTCTGAAGAATATGCCATGAATCTGGGCTATTTGGGTAAAGGAATGTGCT 606
Qy 1500 AGCAAAAAAACAATTTAGAACCAAGTAGAAGCGGGCGCATTTGTTTAAATTCATGAA 1559
Db 607 TATGAACCTTCCCTGTGTCGATCAATCGAAGCTTGGAGCCATTTGGCTTTAAATCCAGAA 666
Qy 1560 GACTGGGGCAACAACAAGTGGCATCGATCACTGCTTGGAGCGTGGCAGATGAATAGAT 1619
Db 667 GACTGGGGTAGCACACCTGCGAGCCATCTACATTTGCTTGAATGTGGCTGCAAAATAGAT 726
Qy 1620 GTGCAAGTTTGTATCCACAGGATACAGTCAATGAGCAGGTTATGTAGATGACACCTTA 1679
Db 727 GTGCAAGTGGCTATCCACACCGATACCTTGAATGAAGCGGGCTGTGTGGAAGACACTTTG 786
Qy 1680 AATGCAATGAAGCGGGCGGCATCCATGCTTACCACATTCAGGAGCGGGTGGAGGACAC 1739
Db 787 CAAGCCATTTGCTGGGCGCATATCCACACTTTCACACTGAAGTGTGTGGCGGGCAC 846
Qy 1740 TCACCTGATGTTATCACCATTGGCAGGAGCTCAATATTCTACCTCTCCACACCCCCC 1799
Db 847 GCTCCGATGTCAATTAAGATGTCTGGCGAATTTAAACATCTCCAGCTTCTACCAACCCC 906
Qy 1800 ACTATTCCTATACCATTAATACGGTTGCGAACAACATTAGACATGCTCATGACATGCCAC 1859
Db 907 ACCATTCTTTTACCGTGAATACAGAAGCCGAAACACATGGACATGTTGATGGTGGCCAC 966
Qy 1860 CACCTAGACAAACGATCCCGGAGATTTACAAATTTCTCAAAGCGGTATCCGCCCGCGC 1919
Db 967 CACTTGGATAAAAACATCAAGAAGATGTCCAGTTTGTGATTTCTAGGATTCGCGCCCAA 1026
Qy 1920 TCTATCCGGCTGAAGATGTCTCCATGATATGGGTGTGATTCGGGATGACAAAGCTCGGAT 1979
Db 1027 ACCATCCGGCTGAGGACAACTCCAGATATGGGATTTCTCTATCACCAGCTCTGAC 1086
Qy 1980 TCGAAGCAATGGGGCTGAGGCGAAGTGAATCTCTGAACTTGGCAGACTGCGGATAAG 2039
Db 1087 TCCCAAGCGATGGGCGGTGATGGGAGGTATCACCGCATCTTGGCAACAGCGGACAAA 1146
Qy 2040 AATAAAAAGATTTGGTTAGCTTCTGAGATGCGAAGATAAGTAATTTCCGCAAT 2099
Db 1147 AACAAAAAGATTTGGTTCGCTGAGGAAAAAGGCGATATATGACAACTTCGCAATC 1206
Qy 2100 AAGCGCTACATCTCCAAATACATATCAACCCCGCTTTGACCCAGCGGTGAGCGAGTAT 2159
Db 1207 AAGCGCTACATTTCCAAATACACCATCAACCCCGCTATTCACACGGCATTTCTGAATAT 1266
Qy 2160 ATCGGCTCTGGAAGAGGGCAAGATCGCGAATTTGGTGGTGTGGAATCTTCGCTTTTTT 2219
Db 1267 GTCGGCTCTGTAGAAGTGGGCAAAATTCGCCGATTTGGTGTCTTTGGAGTCTTCGCTTTT 1326

Qy 2220 GGCGTAAAAACCCAAATTCGTGATCAAAAGCGGTATGTTGTCTTCTCTGAAATGGGCCAT 2279
Db 1327 GGCAATTAACCCCAACATGATCATCAAGGGGGATTATCGACATTTCTCAAAATGGCGCAT 1386
Qy 2280 TCTAAAGCGTCTGTGCCACTCCCACTCCCGTTTATTAACCGGAAATGTTTGGGCATCAC 2339
Db 1387 GCCAATGCCCTCTATCCCCACTCCCAACCCCGTATTAACCGGAAATGTTTGGGCCCAT 1446
Qy 2340 GGCAAGCGGAAATTTGACACAGCATCACTTTTGTTCCTAAAGTCGCTATGAAAAATGGC 2399
Db 1447 GGTAAAGCCAAATTTGACACCAATATCACTTTTGTATCCCAAGTGGCTTATGACAAAGC 1506
Qy 2400 GTGAAAGAAAAGCTGGGCTTAGAGCGCAAGTTCTACGGTCAAAAACTGCGGTAAACATC 2459
Db 1507 ATTAAGAGAGTTGGGCTTGCAGAGAGTGGTTTGGCAGTTAAAAAAGTGGCGCAACATC 1566
Qy 2460 ACCAAGAAAGACTTCAAGTTCAAGCAAAAACGCAAAAATCACCGTCGATCCGAAAAACC 2519
Db 1567 ACCAAAAAAGACCTCAAAATTCACAGTGTACCGCACACATCGAAGTCAATCTCTGAAACC 1626
Qy 2520 TTGAGGTCTTTGTAGATGCAAACTCTGCACCTCTAAACCCACTCGCAAGTGCCTCTA 2579
Db 1627 TACAAGTTAAGTGGATGGCAAGAGTTACTTCCAAAGCAGCGGATAAATCAGCCTA 1686
Qy 2580 GCCAGCGCTACACTTTCTTCTAG 2603
Db 1687 GCACAACTCTACAACTTGTCTTAG 1710

RESULT 8

US-10-476-313-10
; Sequence 10, Application US/10476313
; Publication No. US2004241175A1
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, BRENT
; APPLICANT: BOYLE, JEFFEREY
; APPLICANT: LEW, ANDREW
; TITLE OF INVENTION: Antigen Targeting
; FILE REFERENCE: BDWP-002
; CURRENT APPLICATION NUMBER: US/10/476,313
; CURRENT FILING DATE: 2003-10-24
; PRIOR APPLICATION NUMBER: PRS241
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentln version 3.1
; SEQ ID NO 10
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Helicobacter felis
US-10-476-313-10

Query Match 31.3%; Score 902.4; DB 8; Length 1719;
Best Local Similarity 70.6%; Pred. No. 1.1e-195;
Matches 1203; Conservative 0; Mismatches 501; Indels 0; Gaps 0;

Qy 900 AAAATGAAAAACAAGAAATATGTAATACCTACGGACCCACCAAGGGCGATAAGTGGCG 959
Db 15 AAGATTTCAAGAAAGAAATATGTTCTATGTATGTTCCCACTACCGGGGATCGTGTAGA 74
Qy 960 TTAGAGATACCGATCTTTGGGCAAGATAGAACATGACTATACCACTATGGCGAAGAA 1019
Db 75 CTCGCGACACTGATTTGATCTTAGAAGTGAGCATGATTTGCACCACTTATGTTGAAGAG 134
Qy 1020 CTTAAATTTGGCGGGTAAAAACTATCCGTGAGGGTATGGGTACAGCAATAGCCCTGAT 1079
Db 135 ATCAAAATTTGGGGCGGTAAAACTATCCGTGATGGGATGAGTCAAAACCAATAGCCCTAG 194
Qy 1080 GAAACACCTTAGATTTTAGTCATCACTAACGCGATGATTATCGACTACACCGGATTTAC 1139
Db 195 TCTTATGAATTAGATTTTGGTGTCTACTAACGCCCTCATTTGTGACTATACGGGCATTTAC 254
Qy 1140 AAAGCCGACATTTGGGATTTAAAAACGCGCAAAATTCATGGCATTTGGCAAGGACGAAACAAG 1199

Db 255 |||||AAAGCCGATTTGGGATTAAGACGCAAGATTTGCGGCAATTTGGCAAGGAGCAATAAG 314
Qy 1200 GACATCAAGATGGCGTAAAGCCCTCATATGGTCTGGTGTGGGCAAGAGCACTAGCA 1259
Db 315 GACATGCAAGATGGCGTAGATATATCTTTGGGTAGGTCCTGCTACAGAGGCTTTGGCA 374
Qy 1260 GGGGAAGTATGATTAATACCGCTGGGGGAATCGATTCACACACCACCTTCCTTTCCCA 1319
Db 375 GCTGAGGGCTTGATTGTAACCGCTGGTGGCATCGATACGCATATTCACCTTTATCTCTCC 434
Qy 1320 CAACAATTCCTACCGCTTAGCCATGGGTTTACAACCAATGTTGGAGCGGCAAGGT 1379
Db 435 CAACAATTCCTACTGCTTTGCGAGCGGGTTTACAACCAATGTTGGAGGAGGACAGGA 494
Qy 1380 CTTGTAGATGGCAAGATCGACTACTATCACTCCGGGCAATGGAACCTTGACCCGATG 1439
Db 495 CTTGGGATGGCAAGATGGCAACCATCATCTCCGGAGCGGCTAATCTAAMAAGTATG 554
Qy 1440 TTGCGCGCAGCAGAAGATATTCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCT 1499
Db 555 TTGGGTGACGCGAAGATACGCCATGAATCTAGGCTTTTGGCTAAGGGGATGTCT 614
Qy 1500 AGCAAAAACAATCTGTAGAAACGATAGAACGGGCGGATGGTGTAAATTTGATGAA 1559
Db 615 TAGCAACCTCTTTTACCGCATCAGATTGAAGCAGGGGCGATTGGTTTAAATCCACGAA 674
Qy 1560 GACTGGGCGACACACCAAGTCGATCGATCACTGCTTGAGCGTGCAGATGAATACGAT 1619
Db 675 GACTGGGGAAGCACACTCGAGCTATTCACCATGCTCAATGCTCCGATGAATACGAT 734
Qy 1620 GTGCAAGTTTGTATCCACACCGCATACAGTCAATGAGCGAGGTTATGTAGATGACACCTA 1679
Db 735 GTGCAAGTTGCTATCCACACCGATACCTTTACGAGCGGCTGTGTAGAGACACCTA 794
Qy 1680 AATGCAATGAACGGGCGCCATTCATGCTACCAATGAGGAGCGGTGGAGACAC 1739
Db 795 GAGGCGATTGCGGGGCGCACCATCCATCCTCCACACTGAAGGGGCTGGGGTGGACAC 854
Qy 1740 TCACCTGATGTTATCACCATGCGGCGAGCTCAATATTCACCTCCCTCCACACCC 1799
Db 855 GTCCAGATGTTATCAAAATGGCAGGGAAATTAACAATTCACCCGCTCTACTAACCG 914
Qy 1800 ACTATTCCCTATACCAATTAATACGTTGAGAACACTTACATGCTCATGATCCAC 1859
Db 915 ACATTCCTTTTACCAAAAACACTGAAGCGGACCATGACATGATTAATGTGTGCCAC 974
Qy 1860 CACTAGACAAACGCAATCCGCGAGGATTTACAATTTTCTCAAGCGGTATCCGCCCGG 1919
Db 975 CACTTGGATAAAGTATCAAGGAAGATGTCAGTTTGGCGATTTCGAGGATTCGCCCCAA 1034
Qy 1920 TCTATCGCGGCTGAAGATGTCTCATGATATGGGTGTATGCGGATGACAGCTCGGAT 1979
Db 1035 ACTATCGCGGCTGAAGACCAACTCCATGATGATGGGATCTTTTCTATCACCAGCTCCGAC 1094
Qy 1980 TCGCAAGCAATGGGGCTGAGCGAGGATTCCTCGAACTGGCAGACTCGGATAG 2039
Db 1095 TCTCAGGCTATGGGACGCGTAGCGAGGATGATCACACGCACTTGGCAGACAGACAAA 1154
Qy 2040 AATAAAAAAGAAATTTGGTAAGCTTCTGTAAGATGGCAAGATAACGATAATTTCCGCATT 2099
Db 1155 AACAATAAGATTTGGCGCTTGAAGAGGAAAGGCGATAAACGAACTTCGCATC 1214
Qy 2100 AAGCGTACATCTCATAATACATATCAACCCGCTTTGACCCACCGGCTGAGCGAGTAT 2159
Db 1215 AAACGCTACATCTCTAAATACACCAATCAACCCCGCATCGCGATCGGATTTCTGACTAT 1274
Qy 2160 ATCGGCTCTGTGGAAGGCGAGATCGCGACTTGGTGGTGGATCCTGCTTTTT 2219
Db 1275 GTGGGCTCTGTGGAAGTGGGCAATACGCGGACCTCGCTTTGGAGTCCGGCTTCTTT 1334
Qy 2220 GCGGTAACCAACCAATCTGTATCAAGGCGGTATGGTGTCTCTCTCAAAATGGCGGAT 2279

Db 1335 GGCATTAAGCCCAATATGATTAATTAAGGGCGGATTTATTGCGCTCTCTCAAAATGGGCGAT 1394
Qy 2280 TCTAAACGCGTCTGTGCGCACTCCCAACCGGTTTATTACCGCGAAATGTTTGGGCACTCAC 2339
Db 1395 GCCAATGCGTCTATTCCACCCCTCAGCCGCTCTATTACCGTGAATGTTTGGACACCAT 1454
Qy 2340 GGCAAGCGGAAATTTGACACAGCATCACTTTTGTTCCTCAAGTGGCTATGAAAATGCG 2399
Db 1455 GGGAAAAACAAATTCGACACCAATATCACTTTCTGTGTCCTCAAGGGCTTACAAGCGAGG 1514
Qy 2400 GTGAAGAAAGCTGGGCTTTAGAGCGCAAGTTCTACCGGTCAAAAACCTGCGTAAACATC 2459
Db 1515 ATCAAGAAGAACTAGGGCTAGATGCGGTGATTCAGAGTGAATAAACTGCGCAATATC 1574
Qy 2460 ACCAAGAAAGCACTTCAAGTTCAACGCAAAAACGCAAAAATCACCGTCGATCCGAAAACC 2519
Db 1575 ACTAAAAAGACCTCAAAATTCACGATGTGACCGCACATATTGATGTCAACCTGAAAACC 1634
Qy 2520 TTCGAGTCTTTGTAGATGGCAAACTCTGCACCTCTTAAACCCACTCGCAAGTGCCTCTA 2579
Db 1635 TATAAGGTGAAGTGGATGGCAAGAGGTAACCTCTAAAGCAGCAGATGAATTGAGCCTA 1694
Qy 2580 GCCAGCGCTACACTTTCTTCTAG 2603
Db 1695 GCGCAACTTTATAATTTGTTCTAG 1718

RESULT 9
US-10-335-977-3849
; Sequence 3849, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 3849:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1710 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori

```

;
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...1710
; SEQUENCE DESCRIPTION: SEQ ID NO: 3849:
US-10-335-977-3849

```

Query Match	30.7%	Score 886.4;	DB 7;	Length 1710;
Best Local Similarity	70.0%	Pred. No. 5.1e-192;		
Matches 1193;	Conservative	0;	Mismatches 511;	Indels 0;
Gaps 0;				

Qy	900	AAAATGAAAAACAAGAAATATGTAAATACCTACGGACCCACAAAGCGGATAAAGTCGC	959
Db	7	AAGATTAGCAGAAAAAGAAATATGTTCTTATGTATGGCCTTACCACAGCGGATAAAGTGAGA	66
Qy	960	TTAGGAGATACCGATCTTTGGCGAGAGTAGAATACATGACTATATCCACTATGCGGAGAA	1019
Db	67	TTGGCGGATACAGATTTGATCGCTGAAGTAGAATGACTACACCATTTTATGGCGAAGAG	126
Qy	1020	CTTAAATTTTGGCGGGGTAAAACTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGAT	1079
Db	127	CTTAAATTCGTGGCGGTAAAAACCTTAAGAGAGGCATGAGCCNAATCCAACAACCTTAGC	186
Qy	1080	GA AAAACACCTAGATTTAGTTCATCATAACGCGAGATTAATTCGACTACACGGGATTTAC	1139
Db	187	AAAGAAAGAACTGGATTTTGATCATCACATAAGCGCTTTAATCGTGGATTACACCGGTATTTAT	246
Qy	1140	AAAGCCCAACATTCGGATTTAA AAAACGGCAAAATCCATGGCATTCGCAAGGCAGAGAAACAAG	1199
Db	247	AAAGCGGATATTTGGTATTTAAAGACGGCAAAATCGCTGGCAATGGTAAGGCGGTAAACAA	306
Qy	1200	GACATGCAAGATGGCGTAAGCCCTCATATGGTCTGTGGTGTGGGCA CAGAAGCACTAGCA	1259
Db	307	GACATGCAAGATGGCGTTAAAAACAATCTTAGCGTGGTCTCTACTCTGAAGCGCTAGCT	366
Qy	1260	GGGGAAGGTATGATTTATACCGCTGGGGGAATCGAATTCACACCCACTTCCTTTCTTCCA	1319
Db	367	GGTGAAGGCTTAATCGTAAACGGCTGGTGGTATTGACACACACATCCACTTCATTTCAACC	426
Qy	1320	CAACAATTCCTTACCGCTCTAGCAAATGGCGTTACACCAATGTTTGGAGGGCGCACAGGT	1379
Db	427	CAACAAATCCCTACAGCTTTTGAAGCGGTGTAAACAACCATGATTTGTTGGCGGAACCGGT	486
Qy	1380	CCTGTAGATGGCAGAAATGGACTACTATCACTCCGGGCAAAATGGAAACTTTGCACCGCATG	1439
Db	487	CCTGCTGATGGCACTAATGGCACTACTATCACTCCAGGTAGAGAAACTTTAAATGGATG	546
Qy	1440	TTGCGCGACAGAAAGATATTTCTATGAATGTGGCTTTTTTGGGCAAGGCAATAGTCTCT	1499
Db	547	CTCAGAGCGGTGAAGAATAATTCTATGAACCTTAGGTTTCTTAGCTAAAGGTTAAGCCTTCT	606
Qy	1500	AGCAAAAAACAACTTTGTAGAACAGTAGACGGCGCGGATTTGGTTTTTAAATTCGATGAA	1559
Db	607	AACGACGCGAGCTTAGCGGATCAAAATGAAGCGGTGCGATTTGGCTTTTAAATCCAGAA	666
Qy	1560	GACTGGGGCAACAACCAAGTGGCATCGATCACTGCTTGAGCGTGGCAGATGAATACGAT	1619
Db	667	GACTGGGGCAACACTCCTTTGCGAATCAATCATCGTTAGACGTTTGGGCAAAATACGAT	726
Qy	1620	GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCAGGTTATGTAGATGACACCCCTA	1679
Db	727	GTGCAAGTCGCTATCCACACAGACACTTTTGAATGAAGCCGTTTCGTTGGGAAGACACTATG	786
Qy	1680	AATGCAATGAACGGGGCGGCATCCATGTCCTTACACATTTAGGAGACGGGTGGAGGACAC	1739
Db	787	GCAGCTTATGCGCGACGCACTATGCACATTTTCCACACTGAAGCGCTGGTGGCGGACAC	846
Qy	1740	TCACCTGATGTTATCACCATGGCAGGCGAGCTCAATATTTTACCCTCTCTCCACCAACCCCC	1799
Db	847	GCTCTGACATTTTAAAGTGGCGGTGAACACAAACATCCCTTACCCTGCTTCACTTAAACCC	906
Qy	1800	ACTATTCCTTATACCAATTAATACGGTTTGCAGAAACATTTAGACATGCTCATGACATGCCAC	1859
Db	907	ACTATCCCTTTTCACTGTGTAATACAGAGCGCAACACATGGCAATGCTCATGCTGTGGCAC	966

Qy	1860	CACCTAGACAAACGCATCCCGGAGGATTTACAATTTTCTCAAGACCGTATCCGCGCCCGC	1919
Db	967	CAC TTGGATAAAAGCATATAAGAAGATGTCCAGTTCCGTCGTATTCAAGGATCCGCGCTCAA	1026
Qy	1920	TCATATCGCGCTGAAGATGCTCCATGATATGGGTGTGATCGCGATGACAACGTCGGAT	1979
Db	1027	ACANTTCGGCTGAAAGACATTTGTGCATGACATGGGGATTTTTCTCAATCACAGATTCTGAC	1086
Qy	1980	TCGCAAGCAATGGGCGGTGCGAGCAAGTGAATCTCTCGAACTTTGGCAGACTCGGGATAAG	2039
Db	1087	TCTCAACGCATGGTCTGTGGGTGAAGTTATCACTAGGACTTTGGCAACAGCTGACAAA	1146
Qy	2040	AATAAATAAGAAATTTGGTAAAGCTTCTCGTAAGATGGCAAAAGATTAACGATAATTTCCCGATT	2099
Db	1147	AACAAAAAAGAAATTTGGCCGCTTTGAAAGAAGAAAAAGCGATAACGACAACTTCCAGGATC	1206
Qy	2100	AAGCGCTACATCTCCAAATACATCTACACCCCGCTTTTGACCCACGCGGTGAGCGAGTAT	2159
Db	1207	AAACGCTACTTGTCTAAATACCAATTAACCCAGCGATCCGCTCATGGGATTAGCGAGTAT	1266
Qy	2160	ATCGGCTCTCTGGAAGAGGGCAAGATCCCGACTTGGTGGTGTGGAACTCTGCTCTTTT	2219
Db	1267	GTAGGTTCTGTAGAAGTGGCAAAAGTGCGCTGACTTGGTATTGTGGAGTCCAGCAATCTCTT	1326
Qy	2220	GGCGTAAACCCAAAATCTGTGATCAAAAGGCGGTATGGTGGTCTTCTCTGAATGGGCGAT	2279
Db	1327	GGCGTGAACCCCAACATGATCATCAAAAGGCGGATTCATTCGCTTAAAGCCAAATGGGCGAT	1386
Qy	2280	TCATAACCGGCTGTGCCACTCCCCAACCGGTTTTATTACCGCGAAATGTTTGGGCATCAC	2339
Db	1387	GGCAACGCTTCTATCCCTACCCCAACCCGGTTTTATTACAGAGAAATGTTTCGCTCACCAT	1446
Qy	2340	GGCAAGCGCAAAATTTGACACGAGCATCACTTTTTTGTGTTTCCAAAGTCGCTATGAAAATGGC	2399
Db	1447	GGTAAGCTAAATACGATGCAACATCACTTTTGTGTCTCAAGCGGCTTATGACAAAGGC	1506
Qy	2400	GTGAAGAAAGCTGGGCTTAGAGCGCAAGTTCCTACCGGTCAAAAACCTGCGGTAAACATC	2459
Db	1507	ATTAAGAAGAAATTTAGGGCTTGAAGAACAAGTGTGTGCGGTTAAAAATTTGCAGAAAAATC	1566
Qy	2460	ACCAAGAAAGACTTCAAGTTTCACGCAAAACGGCAAAAATCACCGTCGATCCGAAAACC	2519
Db	1567	ACTAAAAAGACATGCAATTCATATGACATACCGCTCACATTTGAAGTCAATCTCTGAAACT	1626
Qy	2520	TTCCAGGCTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACCTCGCAAGTGCCTCTA	2579
Db	1627	TACCATGTGTTCTGTGGATGGCAAGAAGTCACTTCTAAACGAGTAATAAGTGAGCTTG	1686
Qy	2580	GCCACGCGCTACACTTTCTTCTAG	2603
Db	1687	GCTCAACTCTTTAGCATTTTCTAG	1710

RESULT 10

RESULTS TO
US-09-895-913A-251

US-09-893-913A-231
: Sequence 251. Application US/09895913A; sequence 231, Application
; Patent No. US20020160456A1

; FACEID NO: 052002010
; GENERAL INFORMATION:

APPLICANT: Kleanthous, Harold

APPLICANT: Al-Garawi, Amal

APPLICANT: Miller, Charles

APPLICANT: Tomb, Jean Francois

; APPLICANT: Oomen, Raymond P.

; TITLE OF INVENTION: Identification of Polynucleotides

; TITLE OF INVENTION: Encoding

; TITLE OF INVENTION: Genome

; FILE REFERENCE: 06132/043002

; CURRENT APPLICATION NUMBER: US/09/895,913A

; CURRENT FILING DATE: 2001-06-29

; PRIOR APPLICATION NUMBER: US 08/881,227

; PRIOR FILING DATE: 1997-06-24

; NUMBER OF SEQ ID NOS: 368

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 251

; LENGTH: 1815

; TYPE: DNA

; ORGANISM: Helicobacter pylori

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (51) ... (1757)

US-09-895-913A-251

Query Match 30.7%; Score 885.8; DB 3; Length 1815;

Best Local Similarity 70.0%; Pred. No. 7.2e-192;

Matches 1193; Conservative 0; Mismatches 512; Indels 0; Gaps 0;

QY	900	AAAAAGAAAAACAAGAAATATGTAAATACCTACGAGACCCACCAAGGCGATAAAGTGGCG	959
DB	57	AGATTAGCAGAAAGAAATATGTTCTATGTATGGCCCTACTACAGCGCGATAAAGTGAGA	116
QY	960	TTAGGAGATACCGATCTTTTGGGAGAAAGTAGAAACATGATATACCATCTATGGCGAAGAA	1019
DB	117	TTGGGCGATACAGACTTGATCGTGAAGTAGAAACATGACTACACCATTTTATGGCGAAGAG	176
QY	1020	CTTAATTTGGCGGGTAAACTATCCGTGAGGATGGGTACAGAGCAATAGCCCTGAT	1079
DB	177	CTTAATTTGGGCGGGTAAACCCCTGAGAGAGGATGAGGCAATCCAAACACCCTAGC	236
QY	1080	GAACACACCTAGATTAGTTCATCACTAACCGGATGATTATGACTACACCGGATTTAC	1139
DB	237	AAAGAAAGATTGGATCTAATCATCACTAACCGTCTTAATCGTGGATTACACCGGATTTAT	296
QY	1140	AAAGCGGATTTGGGATTTAAAAACGGCAAAATTCATGGCATTTGGCAAGGCGAGAAACAAG	1199
DB	297	AAAGCGGATTTGGTATTAAAGATGCAAAATCGCTGGCATTTGGTAAAGCGGTAACAAA	356
QY	1200	GACATGCAAGATGGCGTAAGCCCTCATATGCTGGGTGGGCAAGAGCACTAGCA	1259
DB	357	GACATGCAAGATGGCGTTAAAAACAATCTTAGCGTAGGTCTGTACTGAAAGCCTTAGCC	416
QY	1260	GGGGAAGGTATCATTTATACCGCTGGGGAATCGATTTCACACACCATTTCTTCTCCA	1319
DB	417	GGTGAAGGTTTGATCGTAATCTGCTGGTGTATGACACACATCCATTCATTTCACCC	476
QY	1320	CAACAATTCCTTACCGCTTAGCCATGGCGTTTACAAACCATGTTTGGAGCGGCAAGGT	1379
DB	477	CAACAATTCCTTACAGCTTTTGCAGCGGTGTAAACAACCATGATTGCTGGCGGAATGCT	536
QY	1380	CCTGTAGATGGCAAGATCGCATCTACTATCTCCGGGCAATGGAATTTGCACCGCATG	1439
DB	537	CCTGTGTATGGCACTAATGCGACTACTATCACTCCAGGCGAAGAAATTTAAAAATGGATG	596
QY	1440	TTGCGGCGACAGAGAGTATTCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCT	1499
DB	597	CTCAGAGCGGCTGAAGAAATTTCTATGAATCTAGGTTTCTTGGCTAAAGGTAAACGCTCT	656
QY	1500	AGCAAAAAACAATTTGTAGAACAAAGTAGAAGCGGCGGATTTGTTTAAATTTGCATGAA	1559
DB	657	AACGAGCGGAGCTTAGCCGATCAATTTGAAGCTGGTGGCATTTGATTTAAATCCAGAA	716
QY	1560	GACTGGGCGACAAACCAAGTCGATCGATCACTGCTTGAGCGGTGGCAGATGAATACGAT	1619
DB	717	GACTGGGCGACCACTCTCTCTGCAATCAATCATGCTTAGATGTTTGCAGACAAATACGAT	776
QY	1620	GTGCAAGTTTGTATCCACACCGATACAGTCATAGGCGAGGTATGTAGATGACACCCTA	1679
DB	777	GTGCAAGTCGCTATCCACACAGACACTTTTGAATGAAGCGGTTTGGCTGGAGACACTATG	836
QY	1680	AATGCAATGAACGGCGGCCCATCCATGCTTACCATATTGAGGAGCGGTTGGAGGACAC	1739
DB	837	GCAGCTATTGGCGGAGCGCATATGACACTTTCCACACTGAAGGTGCTGGCGGCGACAC	896
QY	1740	TCACCTGATTTATCACCATGGCAGCGAGCTCAATATTCTACCTCTCCACCAACCC	1799
DB	897	GCTCCTGATATTATTAAAGTAGCTGGTGAACACAACTTTCTCCCGCTTCCACTAACCC	956

QY	1800	ACTATTCCCTTATACCAATTAATACGGTTGCAGAACACTTAGACATGCTCATGACATGCCAC	1859
DB	957	ACTATCCCTTTTCACTGTGAATACAGAAGCAGAACACATGGACATGCTTATGGTGTGCCAC	1016
QY	1860	CACCTAGACAAACGCATCCCGGAGGATTACAAATTTTCTCAAGCCGATTCGGCCCGGC	1919
DB	1017	CACCTGGATAAAGCATTAAGAAGATGTTTCAGTTCGCTGATTCAGAGATCCGCCCTCAA	1076
QY	1920	TCATCGCGCTCAAGATGTGCTCCATGATATGGGTGTGATCGGATGACAAAGCTCGGAT	1979
DB	1077	ACATTTGGCGCTGAGACACTTTGTCATGACATGGGGATTTTCTCAATCACCAGCTCTGAC	1136
QY	1980	TCGCAAGCAATGGGGCGTGCAGGCGAGTATTCTCTCGAACTTGGCAGACTGCGGATAG	2039
DB	1137	TCTCAAGCTATGGGTGCTGTGGGTGAAGTTTATCACTAGNACTTGGCAACAGCTGACAAA	1196
QY	2040	AATAAAAAAGAAATTTGGTAAGCTTCTGTAAGATGGCAAGATAAGATAATTTCCGCTT	2099
DB	1197	AACAAAAAGAAATTTGGCGCTTTGAAAGAAAGAAAGGCGATTAACGACACTTCAGGATC	1256
QY	2100	AAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACGGGTGAGCGAGTAT	2159
DB	1257	AAAGCTACTTGTCTTAATACCAATTAACCCAGCGATCGCTCATGGGATTAGCGAGTAT	1316
QY	2160	ATCGGCTCTGTGGAAGAGGCAAGATCGCGGACTTGGTGGTGTGGAATCTCTGCCCTTTT	2219
DB	1317	GTAGGTTCTGTAGAAGTGGGCAAGTGGCTGACTTGGTATTGTGGAGTCCCGCATTTCTT	1376
QY	2220	GGCGTAAAAACCCAAATTCGTGATCAAAAGCGGATATGGTGGTCTTCTCTGAAATGGGCGAT	2279
DB	1377	GGCGTAAAAACCCCAACATGATCATCAAGGCGGGTTCTATGCGTTGAGTCAAAATGGGTGAC	1436
QY	2280	TCATAACGCTCTGTGCGCCACTCCCAACCGGTTTATTACCGCGAAATGTTTGGGCGATCAC	2339
DB	1437	GCAACGCTTCTATTCCTTCCCTACCCCAACCCAGTTTATTACAGAGAAATGTTGCTCATCAT	1496
QY	2340	GGCAAGCGGAAATTTGACACAGCATCATCTTTTGTTCCTTCCAAAGTCGCTATGAAATGGC	2399
DB	1497	GGTAAGCCAAATACGATGCAAAACATCATCTTTTGTGTCTCAAGCGGCTTATGACAAAGC	1556
QY	2400	GTCAAAAGAAAGCTGGGCTTAGAGCGCAAGTCTACCGGTCAAAAACGCGGTAAATC	2459
DB	1557	ATTAAGAGAAATTAGGGCTTGAAGAGACAAAGTGTGCGGTAAAAAATTCAGAAACATC	1616
QY	2460	ACCAAGAAAGACTTCAAGTTCAAGCAAAAACGCAAAATCACCGTCGATCCGAAACC	2519
DB	1617	ACTAAAAAGACATGCAATTTCAACGACACTACCGCTCACATTGAAGTCAATCTCGAAACT	1676
QY	2520	TTGAGGTCTTTGTAGATGCAAACTCTGCACTCTTAAACCCACTCGCAAGTGCCTCTA	2579
DB	1677	TACATGTGTTCTGGATGGCAAGAAAGTAACCTTCTAAACGAGCAATAAAGTGAGCTTG	1736
QY	2580	GCCGAGCGCTACACTTTTCTTAGG 2604	
DB	1737	GCGCAACTCTTTAGCATTTTCTAGG 1761	

RESULT 11

US-10-282-122A-22427

; Sequence 22427, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22427
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Helicobacter pylori
US-10-282-122A-22427

Query Match 30.7%; Score 884.8; DB 7; Length 1710;
Best Local Similarity 70.0%; Pred. No. 1.2e-191;
Matches 1192; Conservative 0; Mismatches 512; Indels 0; Gaps 0;

QY 900 AAAAAAGAAACAGAAATATGTAATACCTACGGACCCACCAAGCGGATAAAGTGCGC 959
DB 7 AGATTAGCAAGAAAGAAATATGTTCTATGATGGCCCTACTACGCGGATAAAGTGAGA 66

QY 960 TTAGGAGATACCGGATCTTTGGGAGAGATGAGAACATGACTATACCACTATGGCGAAGAA 1019
DB TTGGCGGATACAGACTTGATCGCTGAAGTAGAATGACTATACCATTTATGGCGAAGAG 126

QY 1020 CTTAAATTTGGCGGGGTAAGAACTATCCGTAAGGATGATGGGTGAGAGCAATAGCCCTGAT 1079
DB CTTAAATTTGGGTGGCGGTAAGAAACCTGAGAGAGGCAATGAGCAATCCAAACCCCTAGC 186

QY 1080 GAAAAACCCCTAGATTAGTTCATCACTAACCGGATGATTACGACTACACCGGGATTTC 1139
DB AAAGAGAAATTTGATCTAAATCATCACTAAACGCTTTAATCGTGGATTACACCGGATTTAT 246

QY 1140 AAAGCCGACATTTGGGATTAAGAAACCGCAAAATCCATGGCATTTGGCAAGCGGAAACAAG 1199
DB AAAGCGGATATTGGTATTAAGATGGCAAAATCGTGGCATTTGGTAAAGCGGTACAAA 306

QY 1200 GACATGCAAGATGCGGTAAGCCCTCATATGGTGTGGGTGTGGGCAAGAGCACTAGCA 1259
DB GACATGCAAGATGCGGTTAAAGAAACAAATCTTAGCGTAGGTCTGTACTGAAGCCTTAGCC 366

QY 1260 GGGGAGGATGATTATTCGCTGGGGGATCGATTACACACCCATTCCTTTCTCCA 1319
DB GGTGAAGGTTTGTATCGTAATCTGCTGGGTGTTGACACACATCCATCTCATTTACCC 426

QY 1320 CAACAATTTCCCTACCGCTCTAGCCAAATGGCGTTTACAAACCATGTTTGGAGCGGACAGGT 1379
DB CAACAATTTCCCTACAGCTTTTGGAGCGGTGTAAACCAATGATTGTTGGCGGAACTGGT 486

QY 1380 CCTGTAGATGCGGAATGCGACTATCATCTCCGGGCAAAATGGAATTTGCAACCGCATG 1439
DB CTTTATGATGCGGAATTTCAACGCACTACCGCTCATTTGAAGTCAATCTCTGAAACT 1626

DB 487 CCTGCTGATGGCACTAATGCGACTACTATCACTCCAGGAGAGAAATTTAAATGGATG 546
QY 1440 TTGCGCGCAGCAGAGAGATTTCTATGAATGTGGGCTTTTGGGCAAGCAATAGCTCT 1499
DB 547 CTCAGAGCGGCTGAAGAATATTCTATGAATCTAGGTTTCTGGCTAAAGGTAAAGCTTCT 606
QY 1500 AGCAAAAAACAATTTGTAGAACAAAGTAGAACGGGCGGATTTGGTTTAAATTCATGAA 1559
DB 607 AACGACGCGAGCTTAGCCGATCAAAATTTGAAGCTGGTGGATTTGGCTTTAAATTCACGAA 666
QY 1560 GACTGGGCGACAAACCAAGTCGATCGATCACTGCTTGAGCGTGGCAGATGAATACGAT 1619
DB 667 GACTGGGCGACCACTCTCTCTGCAATCAATCATGCTTAGATGTTGACAGACAAATACGAT 726
QY 1620 GTGCAAGTTTGTATCCACACCGGATACAGTCAATGAGGAGGATTTAGTAGACACCCCTA 1679
DB 727 GTGCAAGTCGCTATCCACACAGACACTTTGAATGAAGCGGTTGCGTGGGAAGACACTATG 786
QY 1680 AATGCAATGAACGGGCGGCGCATCCATCCATACCAATTTGAGGAGGAGCGGTGGAGGACAC 1739
DB 787 GCAGCTATTTCGCGGACGCACTATGCACTTTCACACTGAAGGTGCTGGCGGCGGACAC 846
QY 1740 TCACCTGATGTTATCAACATGGCAGGAGCTCAATATTCTACCTCTCTCCACACACCCCC 1799
DB 847 GCTCTGATATTAATAAGTAGCTGGTGAACACAACTTTCCCGCTTCACCTAAACCC 906
QY 1800 ACTATTCCCTATACCAATTAATACGTTTGCAAGACACTTAGACATGCTCATGACATGCCAC 1859
DB 907 ACTATCCCTTTCACTGTGATACAGAAAGCAGACACATGAGACATGCTTATGTTGGTGGCCAC 966
QY 1860 CACTAGACAAACGCACTCCCGGAGGATTTCAATTTTCTCAAGCCGCTATCCGCCCGCC 1919
DB 967 CACTTGGATAAAAAGCATTAAGAAGATGTTCACTGCTGATTTCAAGGATCCGCCCTCAA 1026
QY 1920 TCTATCGCGCTCAAGATGCTCCATGATATGCTGATCGGATGACAAAGCTCGGAT 1979
DB 1027 ACCATTGCGGCTGAAGACACTTTGCTGATGATGGGATTTTCTCAATCACCAGCTCTGAC 1086
QY 1980 TCGCAAGCAATGGGCGGTGAGGCGAAGTATTTCTCGAACTTGGCAGACTTGGGATAG 2039
DB 1087 TCTCAAGCTATGGTCTGTTGGGTGAAGTTATCACTAGAACTTGGCAAAACAGCTGACAAA 1146
QY 2040 AATAAAAAAGAAATTTGGTAAAGCTTCTGAGATGGCAAGATAAGATAATTTTCGCAATT 2099
DB 1147 AACAAAAAGAAATTTGGCGCTTGAAGAGAAAAAGCGGATAACGAACTTTCAGGATC 1206
QY 2100 AAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCAGCGGTGAGGAGTAT 2159
DB 1207 AAACGCTACTTGTCTAAATACACCAITTAACCCAGGATCGCTCATGGGATTAGCGGAT 1266
QY 2160 ATCGGCTCTGTGAAGAGGCAAGATCGCGACTTTGGTGGTGGGATTCCTGCGCTTTTTT 2219
DB 1267 GTAGTTCTGTAGAGTGGGCAAGTGGCTGACTTTGGTATTTGTGGAGTCCCGCATTTCTT 1326
QY 2220 GGCGTAAAAACCAAAATTCGTGATCAAAAGCGGATAGTGGTCTTCTCTGAAATGGGCGAT 2279
DB 1327 GGCGTAAAAACCAACATGATCATCAAAAGCGGTTTCAATGCGCTTGAATTCAGTCAAAATGGGTGAC 1386
QY 2280 TCTAAACGCGCTGTGCGCACTCCCAACCGGTTTATTACCGCGAAATGCTTTGGGCGATCAC 2339
DB 1387 GCGAAACGCTTCTATCCCTACCCCAACCAAGTTTATTACAGAGAAATGTTTCGCTCATCAT 1446
QY 2340 GCGAAGCGGAAATTTGACACCAAGCATCACTTTTGTTCCTCAAGTCCGCTATGAAATGGC 2399
DB 1447 GGTAAAGCAATACGATGCAAAATCATCTTTGTGCTCAAGCGGCTTATGACAAAGGC 1506
QY 2400 GTGAAGAAAAAGCTGGGCTTAGAGCGCAAGTTCTAAGCGTCAAAAACTGCGGTAAACATC 2459
DB 1507 ATTAAGAAGAAATTTAGGCTTGAAGAGCAAGTGTTCGCGTAAAAAATGAGAAACATC 1566
QY 2460 ACCAAGAAAGACTTCAAGTTCAACGCAAAACGCGGAAAAATCAACGTCGATCCGAAAC 2519
DB 1567 ACTAAAAAGACATGCAATTTCAACGCACTACCGCTCATTTGAAGTCAATCTCTGAAACT 1626

RESULT 14
US-09-402-100-1
; Sequence 1, Application US/09402100
; Patent No. US20010019834A1
; GENERAL INFORMATION: Daewoong Pharmaceutical Co, LTD
; APPLICANT: Kim, Byung-O
; APPLICANT: Shin, Sung-Seup
; APPLICANT: Yu, Young-Hyo
; APPLICANT: Park, Myung-Hwan
; APPLICANT: Choi, Deok-Joon
; APPLICANT: Jung, Hyung-Jin
; TITLE OF INVENTION: Recombinant Microorganisms Expressing Antigenic Proteins of Helicobacter pylori
; FILE REFERENCE: 0136/06140
; CURRENT APPLICATION NUMBER: US/09/402,100
; EARLIER FILING DATE: 1999-09-27
; EARLIER APPLICATION NUMBER: KR 97-11950
; EARLIER FILING DATE: 1997-03-31
; EARLIER APPLICATION NUMBER: KR 97-11951
; EARLIER FILING DATE: 1997-03-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2385
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1..)
; OTHER INFORMATION: Recombinant DNA
US-09-402-100-1

Query Match 29.7%; Score 855.6; DB 3; Length 2385;
Best Local Similarity 69.8%; Pred. No. 6.8e-185;
Matches 1169; Conservative 0; Mismatches 504; Indels 1; Gaps 1;

QY 900 AAAAAAACAAGATATGTAATACCTACGAGCCACCAAGGCGGATGAAGTGC 959
DB 7 AGATTAGCAGAAAGAAATATGCTTCTATGATGGCCCTACTACAGCGCATAAAGTGA 66

QY 960 TTAGGAGATACCGATCTTTGGCAGAAAGTAGAACATGACTATACCACTTATGGCGAAGAA 1019
DB 67 TTGGCGATACAGACTTGTATGCTGAAGTAGAACATGACTATACCACTTATGGTGAAGAG 126

QY 1020 CTTAAATTTGGCGCGGTAAACTATCCGTGAGGTATGGGTGAGAGCAATAGCCCTGAT 1079
DB 127 CTTAAATTTGGCGCGGTAAACTATCCGTGAGGTATGGGTGAGAGCAATAGCCCTGAT 186

QY 1080 GAAACACCCCTAGATTTAGTCACTCAACGCGATGATTATGACTACACCGGATTTAC 1139
DB 187 AAGAAAGAACTGGATCTAATCATCACTAAGCTTTTAATCGTGATTACACCGGTATTTAT 246

QY 1140 AAAGCCGACATTTGGGATTTAAACACGCAAAATCCATGGCATTTGGCAAGCGCAGAAACAAG 1199
DB 247 AAAGCGATTTGTTATTTAAAGATGGCAAAATCGCTGGCATTTGGTAAAGCGGTAAACAA 306

QY 1200 GACATGCAAGATGGCGTAAAGCCCTCATATGGTCTGGGGTGTGGGCAAGAACACTAGCA 1259
DB 307 GACACGCAAGATGGCGTTAAAAACAATCTTAGCGTGGGTCTGCTACTGAAGCCTTAGCC 366

QY 1260 GGGGAGGTATGATTTATACCGCTGGGGAATCGATTACACACCCACTTCTTCTTCCA 1319
DB 367 GGTGAAGGTTGATTTGTAATCTGCTGGTGTATTTGACACACATCCACTTCACTCTCCCC 426

QY 1320 CAACAAATCCCTACCGCTCTAGCCATGGCGTTACAAACCATTTTGGAGCGGCGCAGGT 1379
DB 427 CAACAAATCCCTACAGCTTTTGAAGCGGTGTAAACACCATGATTTGGTGGCGGAAGTGC 486

QY 1380 CTTGTAGATGGCAGATGCGATCTACTATCACTCCGGGCAATGGAATTTGACCGCATG 1439
DB 487 CTTGTGTAGTGGCACTAAACGCAACCACTATCACTCCAGGTGAGAAATTTAAAAATTCATG 546

QY 1440 TTGGCGCAGCAGAGAGTATTTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCT 1499
DB 547 CTGAGAGCGGCTGAAGAAATATCTATGAATTTGGTTTCTTGGCTAAAGGTAACGCTTCT 606

QY 1500 AGCAAAAAAACAATTTGTAGAACAGTAGAGCGGCGGATTTGGTTTAAATTTGATGAA 1559
DB 607 AACGATGCAAGCTTAGCCGATCAAAATTTGAAGCTGTGCGATTGGCCTTTAAATTCACGAA 666

QY 1560 GACTGGGGCAACAACCAAGTGGGATCGATCACTGCTTTGAGCGTGGCAGATGAATACGAT 1619
DB 667 GACTGGGGCACCACTCTCTCTGCAATCAATCATGCGTTAGATTTGGGACAAATACGAT 726

QY 1620 GTCAAGTTTGTATCCACACCGATACAGTCAATAGGACGATTTAGTAGAGACACCCCTA 1679
DB 727 GTCAAGTTCGCTATCCACACAGACACTTTGAATGAAGCCGTTGCGTGGGAGACACTATG 786

QY 1680 AATGCAATGAACGGGCGCGCATCCATGCCATACCAATTTAGGAGCGGGTGGAGGACAC 1739
DB 787 GCAGCTATTGCGCGGACGCACTATGCACACTTACCACTGAAGCGCTGGCGGCGGACAC 846

QY 1740 TCACCTGATGTTATCCATGCGGAGCGAGCTCAATTTCTACCTCTCTCCACACCCCTC 1799
DB 847 GCTCTGATATTTAAAGTGGCGGTGAACACACATCTTACCCTGCTTCCACTAACCCCT 906

QY 1800 ACTATTCCCTATACCATTTAATACGGTTTGCAAGACACTTAGACATGCTCATGACATGCC 1859
DB 907 ACTATCCCTTTACCGGTGAATACAGAAAGCCGAAACACATGACATGCTTATGGTGTGCCAC 966

QY 1860 CACTAGACAAACGATCCGCGAGGATTTACATTTTCTCAAGCCGATATCCGCCCGGC 1919
DB 967 CACTTGGATAAAAGCATTTAAGAAAGATGTCAGTTGCTGATTTCAAGGATTTGCGCCTCAA 1026

QY 1920 TCTATCGCGCTGAAGATGTCCTCATGATATGGTGTGATCGGATGACAGATCGGAT 1979
DB 1027 ACCATTGCGGCTGAAGACACTTTGCGATGACATGGGATTTTCTCAATCCTAGTTCTGAC 1086

QY 1980 TCCCAAGCAATGGGGCGTGCAGCGCAAGTATTCCTCGAACTTGGCAGACTGCGGATAAG 2039
DB 1087 TCTCAAGCGATGGCGCGTGTGGGTGAAGTTATCACTAGAACTTGGCAAAACAGCTGACAA 1146

QY 2040 AATAAAAAAGAAATTTGGTAAAGCTTCTGAAAGATGGGCAAGATACGATTAATTTCCGCAAT 2099
DB 1147 AATAAAAAAGAAATTTGGCGCTTTGAAAGAAAGAAAGGCGATACGACAACTTCAGGATC 1206

QY 2100 AAGCGCTACATCTCCAAATACACTATCAACCCGCTTTGACCCACGCGGTGAGCGATAT 2159
DB 1207 AAACGCTACTTGTCTAAATACCAATTAACCCAGCGATCGCTCATGGGATTTAGCGAGTAT 1266

QY 2160 ATCGGCTCTGTGAAGAGGCAAGATCGCCGACTTTGGTGTGTGGAATCTTGCCTTTTTT 2219
DB 1267 GTCGGTCTGTAGAAGTGGCAAGTGGCTGACTTTGGTATTTGGAGTCCCGCATTTCTTT 1326

QY 2220 GCGTAAAAACCCAAATTCGTGATCAAGGCGGTATGGTGTCTTCTCTGAAATGGGCGAT 2279
DB 1327 GGTGTGAACCCCAACATGATCATCAAGGCGGTTCATCGATTTAGTCAAAATGGGTGAT 1386

QY 2280 TCTAAACGCTCTGTGCCACTCCCCAACCGTTTATACCGGAAATTTTGGGCACTAC 2339
DB 1387 GCGAACGCTTCTATCCCTACCCCAACCAAGTTTATTTACAGAGAAATTTGTCGCTCATCAT 1446

QY 2340 GGCAAGCGGAAATTTGACACCCAGCATCACTTTTGTTCCTCAAGTCGCTATGAAAAATGGC 2399
DB 1447 GGTAAAGCTTAATACGATGCAAAACATCACTTTGTCTCAAGCGGCTTATGACAAAGGC 1506

QY 2400 GTGAAAGAAAGTGGGCTTAGAGCGCCAGTTCTACCGGTCAAAAACTGCGGTAACATC 2459
DB 1507 ATTAAGAGAAATTTAGGCTTTGAAAGACAAAGTGTGCGCGTAAAAAATTTGAGAAATATC 1566

QY 2460 ACCAAGAAAGACTTCAAGTTCAACGACAAAAACGCAAAAAATCACCGTTCGATCCGAAACC 2519
DB 1567 ACTAAAAAGACATGCAATTTCAACGACACTACCGCTCACATTTGAAGTCAATTTCTGAAACT 1626

```

Qy      2520  TTGAGGTCCTTTAGATGCAAACTCTGACCTCTAAACCCACCTCGCAAGTG 2573
Db      1627  TACCATGTGTCTGGATGGCAAA-GAAGTAACCTCTAAACCCAGCCATAAAGTG 1679

RESULT 15
US-10-282-122A-32950
; Sequence 32950, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32950
; LENGTH: 2341
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-10-282-122A-32950

Query Match      25.5%; Score 734.6; DB 7; Length 2341;
Best Local Similarity 58.6%; Pred. No. 3.4e-157;
Matches 1408; Conservative 0; Mismatches 929; Indels 56; Gaps 5;

Qy      207  TGAACATCACACCAAGAGCAAGAAAGTTCTTTGTATATATATGCGGCGCAAGTGGCTA 266
Db      2    TGGAAATTAACCAAGAGAAAGATAATTACTGCTTTTACTGCGGCCCTTGTTCGAG 61

Qy      267  GAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGAGACCATTCCTTACATTAGTG 326
Db      62  AAAGACGTTTACGTAAAGGGATTAAAACTTAATTACCTTGAAGCTGTGCGCTTGTATTAGCT 121

Qy      327  CCCATATTATGACGAGCGCGCGTGGAAAAAACCCTTTGCCAGCTTATGGAAGAGT 386
Db      122  GCGCATTTATGGAAGGCCACGAGGGG--GAAACAGTTTCTCAATTAATGATGAGAG 178

Qy      387  GCATGTCATTTTGTAAAAAAGATGAAGTAATGCCCGGGTGGGTAAATATGGTTCCCGATC 446

```

QY	1524	GTAGAACGGCGCGCGATTGGTTTTAAATTCATGAAGACTGGGGCACAACACCAAGTGGC	1583
Db	1259	ATAACAGCGGGTCTATAGGCTTTAAAAATACATGAAGACTGGGGGCAACGCCAATGGCA	1318
QY	1584	ATCGATCACTGCTTGAGCGTGGCAGATGAATAGCATGTGCAAGTTTGTATCCACCGAT	1643
Db	1319	ATTCAAAATTGGCTTAATGTCGCGGATGAAATGGATGTACAGGTGGCTATTCACTCTGAC	1378
QY	1644	ACAGTCAATGAGCGAGGTTATGTAGATGACACCTAAATGCAATGAACGGGGCGGCATC	1703
Db	1379	ACCTTAAATGAAGGTGGTTTTATGAAGAGACAGTAAAGCCATTGCCGTCGAGTGATC	1438
QY	1704	CATGCCCTACCAATTTGAGGAGCGGGTGGAGACACTCACCTGATGTTATCACCATGGCA	1763
Db	1439	CATGTTATCCATACCGAAGCGCGAGGTGGTGTCTATGCCCTGATGTGATCAAGTCGGTA	1498
QY	1764	GGCGAGCTCAATATTTCTACCTCTCCACACCCCACTATTCCCTTATACCAATTAATAG	1823
Db	1499	GGAGAGCCCAATATTTTACCTGCTCAACCAACCCCAACGATGCTTTATACCAATTAATAC	1558
QY	1824	GTTCAGAACACTTAGACATGCTCATGACATGCCACCATAGACAAACGATCCGCGAG	1883
Db	1559	GTGGACGAGCATCTTGATATGTTGATGGTCTGTCTCATCTCGATCCCTCTATTCTCTGAA	1618
QY	1884	GATTTTACAATTTTCTCAAGCGGTATCCGCCCGGCTCTATCGCGCTGAAAGATGTGCTC	1943
Db	1619	GATGTGGCATTTGCTGNACTCTGTAATCTGTTCGTGGGAACCATTTGCTGCAGAAATATCTTA	1678
QY	1944	CATGATATGGGTGTGATCCGATGACAAAGCTCGGATTCGCAAGCAATGGGGGTGCAGGC	2003
Db	1679	CATGATATGGGGCAATTTCCGGTGTGTCGTGAGACTCACAGCCATGGGACGAGTCGGA	1738
QY	2004	GAAATGATTCCTCGAATTTGGCAGACTGCGGATAGAAATAAAAGAAATTTGGTAAGCTT	2063
Db	1739	GAAATGATTCCTGACACTTTGGCAGTGTGCACATAAAATGAAATGCAACGAGGCACATTA	1798
QY	2064	CTGGAAGATGGCAAGATAAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACT	2123
Db	1799	GCGGTGATAGCGCAGATATGATTAATATCGTATTAACGGTTATATCGTAAATACAG	1858
QY	2124	ATCAACCCCGCTTTGACCCACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCAAG	2183
Db	1859	ATTAACCCCGCACTGGCACATGGTATGCTCATACGGTGGGATCAATAGAAAAGGTAA	1918
QY	2184	ATCGCGCACTGGTGTGTGGAATCTGCTTTTGGCGGTAAACCCCAAAATCGTGATC	2243
Db	1919	CTTGGCGATATCGTGTATGGGATCTGCTTTCTTTGGCGTCAAAACCGGCACCTTATCATA	1978
QY	2244	AAAGCGGTATGGTGTCTTCTGNAATGGGCGATTCTAACCGCTCTGTGCCACTCCC	2303
Db	1979	AAAGTGGTATGGTGTGCTTATGCGCAATGGGGGATTAATGCGGCTATTCAACACCG	2038
QY	2304	CAACGGTTTATTACCGGAAATGTTGGGCATCACGGCAAGCGGAAATTTGACACACAG	2363
Db	2039	CAACCGGTTCAATATGCTCAATGATGATGCTGTTAGGAAAAGCCAAATATCAACGTCG	2098
QY	2364	ATCACTTTTGTTCCAAAGTCCGCTATGAAATGCGGTGAAAGAAAGCTGGGCTTAGAG	2423
Db	2099	ATGATCTTTATGTCAAAAGCGGTATTGAGCGGGAGTGCCAGAAAAATTAGGCTTAAAA	2158
QY	2424	CGCCAAAGTCTACCGGTCAAAAACCTGCGTACATCACCAAGNAAGACTTCAAGTTCAAC	2483
Db	2159	AGCTTAAATGGTGTGTGGAGGGCTGTGCTCATATCACAAAAAGCTTCGATGATCCCAAT	2218
QY	2484	GACAAAACGGCAAAAATCACCGTCAATCGAAAACTTCGAGGTCTTTGTAGATGCCAAA	2543
Db	2219	AACTATGTTCTCATATCGAATTAGATCCAAAACTTACATTGTTAAAGGGATGGTGT	2278
QY	2544	CTCTGCACTCTAAACCCCACTCGAAGTGCCTCTAGCCCGAGCGCTTCTTTCTTAG	2603
Db	2279	CCACTGGTTGTGAGCCAGCGACTGAATTACCGATGGCTCAACGCTATTTCTTATTTAA	2338

QY	2604	GCA	2606
Db	2339	CCA	2341

Search completed: November 29, 2005, 08:04:19
Job time : 2060 secs

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).

Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2005, 00:34:45 ; Search time 563 Seconds
(without alignments)
760.989 Million cell updates/sec

Title: US-09-904-994B-1
Perfect score: 2883
Sequence: 1 rgragattttccarctt.....aaaaagtagaacacagg 2883

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3205263 seqs, 74304013 residues
Total number of hits satisfying chosen parameters: 6410526

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New:*
1: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

*Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	612	21.2	3164	1	US-10-793-626-3828
2	612	21.2	3234	1	US-10-793-626-3853
3	578.6	20.1	1716	1	US-10-793-626-117
4	455	15.8	1612	1	US-10-485-517-26
5	424.4	14.7	3475	1	US-10-793-626-4032
C 6	216.2	7.5	526	1	US-10-834-151-2
C 7	97.4	3.4	309	1	US-10-793-626-121
8	74.2	2.6	336	1	US-10-793-626-115
9	46	1.6	2259	1	US-10-793-626-4078
C 10	40	1.4	3373	1	US-10-793-626-4389
C 11	40	1.4	3927	1	US-10-793-626-3926
C 12	39.8	1.4	3923	1	US-10-793-626-4109
C 13	37.8	1.3	1809	1	US-10-793-626-2361
C 14	37.8	1.3	3591	1	US-10-793-626-4033
C 15	37.2	1.3	2988	1	US-10-793-626-4132
16	37.2	1.3	4210	1	US-10-793-626-3761
17	37	1.3	3362	1	US-10-793-626-3731
C 18	37	1.3	4041	1	US-10-793-626-3689
19	36.2	1.3	738	1	US-10-793-626-2909
20	36.2	1.3	3218	1	US-10-793-626-4229
C 21	36.2	1.3	3997	1	US-10-793-626-4220
22	35.8	1.2	14941	1	US-10-821-234-771
C 23	35.6	1.2	1017	1	US-10-793-626-1003

C 24	35.6	1.2	1439	1	US-10-485-517-41	Sequence 41, Appl
C 25	35.6	1.2	1500	1	US-10-485-517-56	Sequence 56, Appl
C 26	35.6	1.2	3189	1	US-10-793-626-3694	Sequence 3694, Ap
C 27	35.6	1.2	3223	1	US-10-793-626-4222	Sequence 4222, Ap
C 28	35.6	1.2	3285	1	US-10-793-626-4141	Sequence 4141, Ap
29	35.6	1.2	340000	7	US-11-102-978-3	Sequence 3, Appli
30	35.4	1.2	3609	1	US-10-793-626-3919	Sequence 3919, Ap
C 31	35.2	1.2	4248	1	US-10-793-626-3885	Sequence 3885, Ap
C 32	35	1.2	2308	1	US-10-510-386-61	Sequence 61, Appl
C 33	34.6	1.2	3473	1	US-10-793-626-3343	Sequence 3343, Ap
C 34	34.4	1.2	1274	1	US-10-927-641-34	Sequence 34, Appl
C 35	34.4	1.2	1299	1	US-10-793-626-779	Sequence 779, App
36	34.4	1.2	3037	1	US-10-793-626-3904	Sequence 3904, Ap
37	34.4	1.2	3073	1	US-10-793-626-3606	Sequence 3606, Ap
38	34.4	1.2	3119	1	US-10-793-626-4270	Sequence 4270, Ap
39	34.4	1.2	3536	1	US-10-793-626-3957	Sequence 3957, Ap
40	34.4	1.2	3645	1	US-10-793-626-4110	Sequence 4110, Ap
41	34.4	1.2	3854	1	US-10-499-715-5	Sequence 5, Appli
42	34.4	1.2	3932	1	US-10-793-626-3420	Sequence 3420, Ap
43	34.4	1.2	4020	1	US-10-793-626-3434	Sequence 3434, Ap
C 44	34.2	1.2	3473	1	US-10-793-626-3343	Sequence 3343, Ap
C 45	34.2	1.2	3546	1	US-10-793-626-3457	Sequence 3457, Ap

ALIGNMENTS

RESULT 1
US-10-793-626-3828
; Sequence 3828, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3828
; LENGTH: 3164
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3828

Query Match	21.2%	Score 612;	DB 1;	Length 3164;
Best Local Similarity	57.4%	Pred. No. 5.2e-139;	Mismatches 845;	Indels 52;
Matches 1211;	Conservative 0;			Gaps 4;
QY	543	TTGGTTGCGATAAGACATCGAGCTCAATCGAGCGCAAGAGAGTACCGAACTTGAGGTTA	602	
Db	468	TTGTTAAAAATCTAGTAATAGAGTCAATAACATCATCCCGAAACGGTGTGAGTAA	527	
QY	603	CTAATGAAGGGCCTAAATCCTTTGCATGTGGGTAGGTCATTCCACTTTCTTTGAAGCTACA	662	
Db	528	AAAATACAGGGGATAGACTATACAACTAGGTTCACATTTCCACTTTTTCGAAGCAATA	587	
QY	663	AGGCNCTAAATTCGATCGTGAAGGCTATGGCAAGCGCTAGATATTCCTCTTGTCGCA	722	
Db	588	AAGCATTTAGAAATTTGATCGTGAGAAAGCATATGTTAAACATTTGATATTCCTCGAGGAG	647	
QY	723	ACACCTACGATTTGGGCGAGCAAAACCCGAAAGTCAGTTGATTCCTCTTGTCGCA	782	
Db	648	CTCAGTCAGATTTGAACTCGAGATGAAAAAAGTACAACTTCTCGAATATTCGAC	707	
QY	783	GTAATAAAGTGTGGCATGAACGGGCTTGTTGAATAAC---	839	
Db	708	GACGTAAATTTATGGATTTCGTTGGTTCGATTCGATGCGGATATTGACGAAGAACCGGTAT	767	

QY 840 AACATAAGCGCTTGACAGGCGAAATCTCAGATTATTCAGTAAGAGACTCCCATG 899
 DB 768 TCCGTCAATGATTCAAATCAAAACGCGCGTTAAAGATGACAGCGCAAGCATG 827
 QY 900 -----AAAAAGAAAAAAGAAATATGTA 925
 DB 828 CGAATAAAAAAGTGTAAATAGCATGATTTAAATGACACAAATCTCAATCACAAAG 887
 QY 926 TACCTAGGAGCCCAAGGCGATTAAGTGGCTTGAGATACCGCATCTTTGGGCGA 985
 DB 888 TCTTTATGACCAACTGAGAGACTCTGTAGATTAAGAGATTCGAACTGTGTCACA 947
 QY 986 AGTAGACATGACTATAACCACTATGCGAGAACTTAATTTGGCGCGGTAAACAT 1045
 DB 948 AGTTGAAAAAGACTATGCAAAATATGAGATGAGACTCTTCGGTGGCGAAAAATCAT 1007
 QY 1046 CCGTAGAGGTATGGGTGCA-----GCAATAGCCCTGATGMAAACACCTTAGATT 1096
 DB 1008 TCGTAGATGTATGGCTCAAAATCTAATGTGACAAAGATGATATAAAATGTAGCCGATT 1067
 QY 1097 AGTCATCACTAACCGGATGATTAATGACTACACCGGATTTACAAAGCGCATTTGGAT 1156
 DB 1068 AGTTTAACTAACCGATTAAATTAATGATTAATGACAAAGATTGTTAAAGCATATCGAAT 1127
 QY 1157 TAAAAACGGCAAAATCCATGCGATTGGCAGAGCAAGAAAGACATGCAAGATGGCGT 1216
 DB 1128 TAAAAATGTATATTTTAAATGATGCTGTAAGCTGMAAACCAAGATTAATGATTAAGT 1187
 QY 1217 AAGCCCTCATATGTGTGTGGGTGTGGGACAGAAAGCACTAGCAGGGGAAGTATGATT 1276
 DB 1188 -----TGACATCATATTGGTGCAACAACTGATTTATGTGTGTGTAAGGTAATTTGT 1241
 QY 1277 TACCGCTGGGGGAATTCGATTACACACCACTTCTCTTCTCACAACAATTCCTAACCG 1336
 DB 1242 TACTGCGGCGGTATCGATACACAGTCACTTCAATCCTGGAACAGCTGAAGTTGC 1301
 QY 1337 TCTAGCCAAATGGCGTTACAACATGTTTGAAGCGGCGCAGGTCTGTAGATGGCAGCA 1396
 DB 1302 ACTTGAAGTGTATTAACAAGCATATCGTGGAGAACTGTGCTTCTGAAAGGTCTAA 1361
 QY 1397 TGGCACTACTATCACTCCGGGCAATGGAATTGCAACCGCATGTTCCGCGCAGCAAGA 1456
 DB 1362 AGCGACTACTGTAAACACAGAGCCTTGGGATTTATCATGCAATGTTAAGAGCAGCAAGA 1421
 QY 1457 GTATTTATGAAATGTGGCTTTTGGGCAAGGCAATAGCTTACCAAAAAACAATTGT 1516
 DB 1422 GATGCTATTAATGTAGATTACTGTAAAGGTCAAGCTGTCAATCACTACCTGCACTAT 1481
 QY 1517 AGAACAAGTAGAAGGGGCGGATGTTGTTTAAATTTGATGAAGACTGGGGCAACAACC 1576
 DB 1482 TGAACAATAATTCATGAGGGGCTATAGGCTTTAAAGTACAGAAATTTGGGGAGCTACACC 1541
 QY 1577 AAGTGCATTCATCACTGCTTGAGCGTGGCAGATGAATACATGTGCAAGTTTGTATCCA 1636
 DB 1542 TTTACGATTAATGATTCATGATTAAGAGTTGCAATGATTAATGTTCAAGTGGCTTTACA 1601
 QY 1637 CACCGATCACTCAATGAGCGAGTTATGTAGATGACACCTTAAATGCAATGACGGCG 1696
 DB 1602 TGCAGATACATTAATGAAGCTGATTTATGMAATACAAATGGCTGTGTAAGATG 1661
 QY 1697 CGGCATTCATGCTTACCACTGAGAGGCGGTGAGAGCACTACCGTGAATGTTATGAC 1756
 DB 1662 TGTATTTGATATGATTCATCACTGAAAGAGCTGTGTGTATGACACTGACTTAATCA 1721
 QY 1757 CATGCGAGCGAGCTCAATATTTCTACCTCTCCACACCCCACTATTCCTATACAT 1816
 DB 1722 ATCAGCTGCATATTCAAACATCTTACCTTCTTCAAAACCAATTAACCTTACCTCA 1781
 QY 1817 TAATACGTTGCAAACTATGACATGTCTCATGACATCCACCACTTACGACAAACGAT 1876
 DB 1782 CAACACTGTATGATGAACATTTAGACATGGTATGATTACTACACCAATCTTAAATGCTTCAAT 1841

QY 1877 CCGGAGGATTTACATTTTCTCAAAAGCCGTATCCGCCCGGCTTATCGCGGCTGAAGA 1936
 DB 1842 ACCAGAGAACAATTGCAATTTGAGATTCTGTATACGTAAAGAAACTATAGACGCAAGA 1901
 QY 1937 TGTGTCCATATATGTTGTGTGATGTGGATGACAAAGCTCGATTGCGAAGCAATGGGGCG 1996
 DB 1902 CGTATTAACAAGATATGCGGTATTTAGTATGTAAGTTCAAGTTTACAAAGCATGGAG 1961
 QY 1997 TGCAGGCGAAGTGAATCTCTCAACTTGGCAGACTCGGATTAAGAAATTAAGAAATTTGG 2056
 DB 1962 TGTGTGTAAGTTGTAACACGACTTGGCAAGTTGACACCGTATGAAAGAAACAGCGG 2021
 QY 2057 TAACTTCTGTAAGATGCGAAAGATACGATATTTCCGACTTAACCGCTACATCTCAA 2116
 DB 2022 ACCATTAGATGTGACTTTGATATATACAGATTAATGTATGTTAAACCTTACATGCAAA 2081
 QY 2117 ATACATTAACACCCCGCTTTGACCCAGCGCTGAGCGAGTATTCGCTGTGTGAAAGA 2176
 DB 2082 ATATACATCAATCTGCAATTAACAGTGAATTTCTGACTATGTTGATCTGTAGAAC 2141
 QY 2177 GGGCAAGATCGCCGACTGTGTGTGGAATCTGCTTTTGGCGTAAACCCAAAT 2236
 DB 2142 GGTAAACTTGGCGATTATGATATGTGGAAACAGAAATTTCTGCGTGCACCAATCC 2201
 QY 2237 CGTGATCAAGCGGTATGTGTCTTCTGAAATGGCGATTTAAACGCTGTGTGCC 2296
 DB 2202 TGTGTTAAAGGTGCAATGATTAATCTAGCAGATTAATGGTATGCTAAATGCTCCATACC 2261
 QY 2297 CACTCCCAACCGGTTTATTAACCGGAAATGTTTGGGATTCACGGCAGCGCAATTTGA 2356
 DB 2262 AACATCAGAGCTTTGAAATATCGCAAAATGTATGTCAATTTGGTGTACATTACACA 2321
 QY 2357 CACGACATCACTTTGTTTCCAAAGTGCCTATGMAAATGCGTGAAGAAAGAAAGCTGGG 2416
 DB 2322 TACTCTATGACTTTTGTGTTCTTAAACCTGCAATATGMAACGATTTATTCGCAACTCA 2381
 QY 2417 CTTAGAGCGCAAGTTCTACACCGGTCAAAATCGCGTAAATCATCACCAAGAAAGCTTCA 2476
 DB 2382 TCTAAGAAAGATGTTTGCAGCGATTAGAAATTTAGAAATTTAACTAAGGCAAGATGAA 2441
 QY 2477 GTTCAACGCAAAACGGCAAAATATCACTGATTCGAAACCTTGGAGTCTTTGTAGA 2536
 DB 2442 AATATATATGCTTACACCTTAAATATGATGATCAACAAATATGAGGTATCGTTGA 2501
 QY 2537 TGGCAAACTGTGACCTCTTAAACCACTGCAAGTGCCTTAGCCGACGCTACCTT 2596
 DB 2502 TGTATATTAATTCACAGATGAAGCAGCAACAGAAATTAACATTAACCAAGATACCTT 2561
 QY 2597 CTTCTAGG 2604
 DB 2562 ATTCTAGG 2569

RESULT 2
 US-10-793-626-3853
 ; Sequence 3853, Application US/10793626
 ; Publication No. US20050255478A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KIMBERLY, WILLIAM JOHN
 ; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
 ; FILE REFERENCE: PU3480US
 ; CURRENT APPLICATION NUMBER: US/10/793,626
 ; PRIOR FILING DATE: 2004-03-04
 ; PRIOR APPLICATION NUMBER: 60/164,258
 ; NUMBER OF SEQ ID NOS: 4472
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 3853
 ; LENGTH: 3234
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic

OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3853

Query Match 21.2%; Score 612; DB 1; Length 3234;
Best Local Similarity 57.4%; Pred. No. 5.3e-139;
Matches 1211; Conservative 0; Mismatches 845; Indels 5

Qy	543	TTGGTTGGGATAAAGACATCGAGCTCAATGACGAGCAAGAAGTAACCGAACTTGAGGTTA	602
Db	537	TTGTTAAAAATACTGAAATAGAAGTCAATAAACATCATCCCGAAACGCTGATTAAGTAA	596
Qy	603	CTAATGAAGGCGCTAAATCCCTTGATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAA	662
Db	597	AAAATACAGGCGATAGACCTATACAAGTAGGTTACATTTCCACTTTTCGAAAGCAATA	656
Qy	663	AGGCACTAAAAATTCGATCGTGAAATAAGCTTATGGCAAAACGCTAGATATTCCTCTGGCA	722
Db	657	AAGCATTAGAATTTGATCGTGAGAAAGCATATGTTAAACATTTTGGATATTCCTGCGAGG	716
Qy	723	ACAGCTACGATTTGGGGCAGCAAAACCGCAAGTCAGTTGATTCCTCTGGTGCGCA	782
Db	717	CTGAGTGAGATTTGAACCTGGAGATGAAAAAAGTACAACTTGTGCAATATTTCTGGAC	776
Qy	783	GTAAAAAGTAGTTGGCATGAACGGGCTTGTGAATAAC---ATCGCGGATGAAGCCCAT	839
Db	777	GACGTAAAAATTTATGGATTCCGTGGTTTAGTCGATGGCGATATTGACGAAGAACGCTAT	836
Qy	840	AACATAAAGCCCTTGACAAGCGGAATCTCACGGATTTTATCAAGTAAAGGAGACTCCCATG	899
Db	837	TCCGTCCAATATGATTCAAATCAAAACCGCCGCTTAAAAACGATGTCAGCGGAGACAAATG	896
Qy	900	-----AAAAATGAAAAAACAAGAAATATGTAAA	925
Db	897	CGAATAAAAAAGTGGTAAATAAGCATGAGTTTAAAAATGACAACTCTCAATACACAG	956
Qy	926	TACCTACGGACCCACAAAGCGGATAAAGTCGGCTTAGGAGATACCGATCTTTTGGGCA	985
Db	957	TCCTTATGGACCAACTGTAGGAGACTCTGTGAGATTTAGGAGATACGAATCTTTTGGCACA	1016
Qy	986	AGTAGAACATGACTATACCACTATGGGGAAGAACTTAAATTTTGGCGGGTAAACATAT	1045
Db	1017	AGTTGAAAAAGACTATGCAATTTATGAGATGAAGCTACTTTTCGGTGGCGGAAATCAAT	1076
Qy	1046	CCGTGAGGTTATGGGTGAGA-----GCAATAGCCCTGATGAAAAACCCCTAGATTT	1096
Db	1077	TCGTGATGTTATGGCTCAAAATCTTAATGTACAAGAGATGATAAAAAATGTAGCCGATTT	1136
Qy	1097	AGTCATCACTAAACGGATGATTCGACTACACCGGATTTTACAAAGCCGACATTTGGAT	1156
Db	1137	AGTTTTAACTAACGCATTAATTTATTTGATTTATGAAGAATTTTAAAGCAGATATCGGAAT	1196
Qy	1157	TAAAAACGGCAAAATCCATGCGATTTGGCAAGCGAGAAACAAGGACATGCAAGATGGCGT	1216
Db	1197	TAAAAATGGTTATATTTTAAAGATCGGTAAAGCTGGAAACCCAGATATAATGGATAACGT	1256
Qy	1217	AAGCCCTCATATGTTGTTGGGTGGGCAAGAAGCACTAGCAGGGGAAGTATGATTAT	1276
Db	1257	-----TGACATCATATGGTGCAACAACATGATATTAATTCCTGCTGAAGGTAAATTTGT	1310
Qy	1277	TACCGCTGGGGGAATCGAATTCACACACCCACTTCTCTTCTCCACAACAATTTCCCTACCGC	1336
Db	1311	TACTGCCGGCGGTATCGATACACACGTCGCACTTCATCAATCTGNAACAAGCTGAAGTTGC	1370
Qy	1337	TCTAGCCAATGGCGTTACAAACCATGTTTGGAGGGCGGACAGGTCCTGTAGATGGACGAA	1396
Db	1371	ACTTGAGAGTGGTATTAACAACGCATATCGGTGGAGGAACTGGTGTCTCTGAAGGTGCTAA	1430
Qy	1397	TGCCACTACTATCACTCGGGCAATGGAACCTTCACCGCATGTTTGGCGCAGCAGAGA	1456
Db	1431	AGCGACTACTGTAAACAAGGACCTTGGCATATTTCAATCGCATGTTTAGAAGCAGCAGAGA	1490
Qy	1457	GTATTTCTATGAATGTGGGCTTTTTTGGGCAAGGCAATAGCTCTTAGCAAAAAACAACCTGT	1516

1491	DB	GATCGCTTAATGATGAGATTTACTGTGTAAGGCTCAAGCTGTCAATCATACTGCACTTAT	1551
1517	QY	AGAAACAAGTAGAAGCGGCGCGAATGTGTTTTAAATTTGCATGAAGACTGGGGCAACAACC	1576
1551	DB	TGAACAAATTCATGCAGCGCTATAGGCTCTTAAAGTAGATGAAGATTTGGGGAGCTACACC	1610
1577	QY	AAGTGGCATCGATCACTGCTTTGAGCTGGGAGATGAATACGATGTGCAAGTTTGTATCCA	1636
1611	DB	TTGAGCATTAAGTCAATGCAATTTAGACGTTGCAGATGAGTTGATGTTCMAGTCGCTTTACA	1670
1637	QY	CACCGATACAGTCAATGAGGCGAGGTATGTAGATGACACCCTAAATGCAATGAACGGGCG	1696
1671	DB	TGCAGATACATTAATGAGCTGGAATTTATGGAAGATACAATGGCTGCTGGAAGATCG	1730
1697	QY	CGCCATCCATGCTTACCACATTTGAGGAGCGGGTGGAGGACACTCACTCGATGTTATCAC	1756
1731	DB	TGTATTGCATATGATCATACTGAAGAGCTGGTGGTGGTCATGCACCTGACTTAATCAA	1790
1757	QY	CATGGCAGGCGAGCTCAATATTCTACCCCTCCTCCACACCCTTATTTCCCTTATACCAT	1816
1791	DB	ATCAGCTGCAATATTCAAACATCTTACCCTCTTCTACAAACCCCAACATTAATCTTACACTCA	1850
1817	QY	TAATACGGTTGCAGAACACTTTAGACATGCTCATGACATGCGACACACCTAGACAAACGCAT	1876
1851	DB	CAACACTGTAGTGAACATTTTAGACATGGTTATGATTACTACACATCTTAATGCTTCAAT	1910
1877	QY	CCGCGAGGATTTACAATTTTTCTCAAAGCCGATATCCGCCCGGCTCTATCGCGCTGAAGA	1936
1911	DB	ACCAGAAGACATTCGCAATTTGCGATATCTCGTATACGTPAAGGAAACTATAGCAGCAGAAGA	1970
1937	QY	TGTGCTCCATGATATGGGTGTGATCGCGATGACAAGCTCGGATTCGGAAGCAATGGGCG	1996
1971	DB	CGTATTCAAAGATATGGGCGCTATTTAGTATGSGTAAGTTTCAGATTCAACAAGCAATGGGACG	2030
1997	QY	TGCAGGCGAAGTGATTCCTCGAATCTTGGCAGACTGCGGATAAGAAATAAAAGAAATTTGG	2056
2031	DB	TGTCGGTGAAGTTGTAAACAGCTACTTGGCAAGTTGCACACCGTATGAAGAGACACGCGG	2090
2057	QY	TAAAGCTTTCCTGAAGATGGCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAA	2116
2091	DB	ACCATTAGATGGTGACTTTTGAATATCAGCATATAATATCGTATTAAAGCTTTACATTTGCAA	2150
2117	QY	ATACATATCAACCCCGCTTTTGACCAACGGCGGTGAGGAGTATATCGGCTCTGTGGAAGA	2176
2151	DB	ATATACAATCAATCCTGCCATTAACATGCTATTTCTGACTATGTTGGATCTGTGAAGC	2210
2177	QY	GGGCAAGATCGCGCACTTGTGTGTGGATCTCGCTTTTTTGGCGTAAACCCCAAAAT	2236
2211	DB	GGGTAAACTTTGCCGATTTAGTAAATGTGGGAACCAGAAATTTCTCGGTGCCAAACCCGATCT	2270
2237	QY	CGTGATCAAAAGGCGGTATGCTGTGCTTCTCTGAAATGGGCGAATCTTAAACGCTCTGTGCC	2296
2271	DB	TGTTGTTTAAAGTTGGCATATTAACCTCAGCAGTAAATGGTGATGCTAAATGGCTCCATACC	2330
2297	QY	CATCTCCCAACCCGTTTATTACCGCGAAATGTGTTTGGGCATCAACGCAAGCGGCAAAATTTGA	2356
2331	DB	AACATCAGACGCTTTGAAATATCGCAAAATGTATGTGTCAAATTTTGGTGGTAAACATTACACA	2390
2357	QY	CACCAAGATCACTTTTGTGTTTCCAAAGTCGCTATGAAATGGCGTGAAGAAAGCTGGG	2416
2391	DB	TACTGCTATGACTTTTGTGTTTCTAACACTGCAATATGAAACGGTATTTTATCGTCAACTCAA	2450
2417	QY	CTTAGCGCGCAAGTTCTACCGGTCAAAAACCTGCCGTAAACATCAACCAAGAAAGACTTCAA	2476
2451	DB	TCTTAAACGAATGGTTCGACAGTTAGAAATATTAAGAAATTTAACTTAAGCGAGATATGAA	2510
2477	QY	GTTTCAACGACAAAACGGCAAAAATTCACCGTCGATTCGAAACCTTCGAGGTCTTTGTAGA	2536
2511	DB	AAATAATAATGCTACACCTTAAATAGATGTAGATCCACAAACATATGAGGTATTCGTTGA	2570
2537	QY	TGGCAAACTCTGCACCTCTTAAACCCACCTCGCAAGTCCTCTAGCCCGCTACACTTT	2596
2571	DB	TGTTAATAAATCACAGTGAAGCAGCAACAGAAATTAACCAATTAACAACAAAGATCTTCTT	2630

QY 2597 CTCTAGG 2604
Db 2631 ATTCTAG 2638

RESULT 3

US-10-793-626-117
; Sequence 117, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 117
; LENGTH: 1716
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-117

Query Match 20.1%; Score 578.6; DB 1; Length 1716;
Best Local Similarity 59.5%; Pred. No. 4.8e-131;
Matches 1021; Conservative 0; Mismatches 679; Indels 15; Gaps 2;

QY 898 TGAATAATCAAAAACAAGATATGTAAATACCTACGACCAACCAAGCCGTAAGTGC 957
Db 8 TTTAAATGACAACTCTCAATACCAAGTCTTTATGACCACTGTAGAGACTCTGTGA 67
QY 958 GCTTAGAGATACCATCTTTGGGCAAGTAGAATAGATACATGATACCACTTATGCGAAG 1017
Db 68 GATTAGAGATACCAACTCTTTGGGCAAGTAGAATAGATACATGATACCACTTATGCGAAG 127
QY 1018 AACTTAAATTTGGCGCGGCTAAACTATCCGTGAGGATATGCGTGA-----GCA 1068
Db 128 AAGCTACTTTCGTGGCGCGAATAATCAATTCGTGATGATGCTCAAAATCCTAATGTGA 187
QY 1069 ATAGCCCTGATGAAACACCTGATTTAGTATATCTATACGCGATGATATGCGATACA 1128
Db 188 CAAGAGATGATAAATATGAGCCGATTTATTTAACTAACGCAATTAATATGATTTATG 247
QY 1129 CCGGATTTACAAAGCCGATTTGGGATTTAAACCGCAAAATCCATGCGATTTGCGAAG 1188
Db 248 ACAAGATTTTAAAGCAAGATATCGAATTTAAATATGTTATTTTAAAGTCTGTAAG 307
QY 1189 CAGAAACAAAGACATGCAAGATGCGTAAAGCCCTCATATGCTGCTGGGTGCGGCAAG 1248
Db 308 CTGAAACCCAGATATATATGATTAAGT-----TGAATCATCATTTGGTGAACAACCTG 361
QY 1249 AAGCATGCGAGGGAAGATATATTTACCGCTGGGGGAATCGATTTACACACCCACT 1308
Db 362 ATATATATGCTGCTGAAGGTAAATTTGTTACTGCGCGGTATCAATACACGCTGCACT 421
QY 1309 TCCCTTCCCAACAATTCCTACCGCTCTAGCCATATGCGGTTTCAACATGTTTGGAG 1368
Db 422 TCATATATCTGAACAAGCTGAAGTTGCACTTGAAGTGTATTTACAGCGCATATCGGTG 481
QY 1369 GCGGCAAGTCTCTGTAGATGCGAAGATGCGATCTACTATCACTCGCGCAATGGAAGT 1428
Db 482 GAGGAATCGTGTCTCTGAAGTGTCTAAAGCACTACTATCAACAGGACCTTGGCATTA 541
QY 1429 TGCACCGATGTTGGCGCGGACAGAAAGATATTTATGATATGTTGGCTTTTGGCGAAG 1488
Db 542 TTCAATGATGTTTGAAGACAGCAAGAGATGCTATTAATGATGAGATTTACTGTTAAG 601

QY 1489 GCAATAGCTCTAGCAAAAAACAATTGTAGAAACAAGTGAAGCGGCGCGATGGTTTTA 1548
Db 602 GTCAAGCTGTATCATATCTGACATTTATGAAACAATTCATGACGCGCTATAGTCTTTA 661
QY 1549 AATTGCATGAAGACTGGGGCAACAACCAAGTGGATGATCACTGCTTGAAGCGGGAG 1608
Db 662 AAGTACATGAAGATTTGGGAGCTACCTTTAGCACTTAAGTATGATTTAGACGTTGAG 721
QY 1609 ATGAATACAGATGTGCAATTTGTATTCACACCGATACAGTCAATGAGGAGGTTATGAG 1668
Db 722 ATGAGTTGATGTTCAATCGCTTTTACATGCAAGATACATTAATGAAGCTGATTTATAG 781
QY 1669 ATGACACCTTAATGCAATGAACGGGCGCGCATCATGCTTACCACTTGAAGGAGCGG 1728
Db 782 AAGATACAAATGCTGCTGTGAAGATCGTGAATTCATATGATATCATCTGAAGAGCTG 841
QY 1729 GTGAGGACATCTACCGTATGTTATACCATGAGGAGGAGGAGTCAATTTACCTGCT 1788
Db 842 GTGGGTGATGACACTTACCTTAATCAATCAGCTGCAATTTCAACATCTTACCTTT 901
QY 1789 CCACACCCCACTATTCCTATACATTTAATACGTTGACAGAACCTTAGACATGCTCA 1848
Db 902 CTACAAACCCAACTATTAACCTTACACTCAACACATGATGATGATTAAGATGATGTTA 961
QY 1849 TGACATGCCACCACTAGCAAAACGATCCGCGAGATTTACATTTTCTCAAGCCGTA 1908
Db 962 TGATTACTCACCATCTTAATGCTTCAATCAGAGACATGCTTTCAGATTCCTGTA 1021
QY 1909 TCCGCCCCGCTCTATCCGCTGTAAGATGCTCATATATGATGATGATGATGATGATG 1968
Db 1022 TACGTAAAGAAATCTATACAGCAAGAGATTTAAAGATATGAGGCGATTTAGTATAG 1081
QY 1969 CAAGCTCGAATTCGCAAGCAATGGGCGTGCAGCGAAGTGAATTCCTGAACTTGGCAGA 2028
Db 1082 TAGATTCAATTCACAGCAATGAGACGCTGCTGATGATGATGATGATGATGATGATG 1141
QY 2029 CTGCGATTAAGATTAATAAAGATTTGGTAAAGCTTCTGAAGATGCAAAAGATTA 2088
Db 1142 TTGCAACCGATTAAGAAAGAACCGGACCTTATGATGATGATGATGATGATGATGAT 1201
QY 2089 AATTCCGATTAAGGCTATCATCTCAATATACATATCAACCCGCTTGAACCCAGCGG 2148
Db 1202 ATATATGATTAATAAGTTATCAATTTGCAAAATATACATCAATCTTCCATTTACATG 1261
QY 2149 TGAAGGATATATGAGCTCTGTGGAAGGAGCAAGATGCGCACTTGGTGTGGAATC 2208
Db 1262 TTTCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1321
QY 2209 CTGCTTTTGGCGTAAACCCCAATCGTATCAAAAGCGGTATGATGATGATGATGATG 2268
Db 1322 CAGAAATTTCTGGTGCACAAACCGAATCTTGTATTAAGGTGATGATGATGATGATG 1381
QY 2269 AAATGGCGATTTAAACCGCTGTGCGCATCTCCCAACCGGTTATTAACCGCAAAATG 2328
Db 1382 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1441
QY 2329 TTGGGATACAGGCAAGGGAATTTGACACAGATCACTTTGTTTCAAAAGTGCCCT 2388
Db 1442 ATGATCAATTTGGTGTAAACATTAACATATGCTATGATGATGATGATGATGATG 1501
QY 2389 ATGAATAATGGCTGAAGAAAGCTGGGCTTGAAGCGCAAGTTTACCGGTCAAAAGT 2448
Db 1502 ATGAATAATGGCTGAAGAAAGCTGGGCTTGAAGCGCAAGTTTACCGGTCAAAAGT 1561
QY 2449 GCGGTATCACTCAAGAAAGCTTCAAGTTCAACGACAAACGCGCAAAATCAACGCTG 2508
Db 1562 TTAGAAATTTAATCTAAGGAGATATGAAGAAATATATATCTACCTTAAATATGATG 1621
QY 2509 ATCCGAAATCTTCAAGGCTTTTGAAGATGCAAACTGCACTCTTAACCAACCTCG 2568
Db 1622 ATCCCAAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1681
QY 2569 AAGTCCCTAGCCCAAGCGCTACACTTTCTTAG 2603

GENERAL INFORMATION:
APPLICANT: KIMBERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PUS480US
CURRENT APPLICATION NUMBER: US/10/793.626
CURRENT FILING DATE: 2004-03-04
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4032
LENGTH: 3475
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4032

Query Match 14.7%; Score 424.4; DB 1; Length 3475;
Best Local Similarity 55.1%; Pred. No. 1.6e-93;
Matches 1010; Conservative 1; Mismatches 762; Indels 59; Gaps 7;

QY 7 ATTTTCACCACTTCAAGCACAATATGATCCTGTGTGGTGGTGTGAATTCACCTTGT 66
DB 1653 AATATTCATCATATCAATTAATTAATTTTACTAAATAATGTTATGTACCGTGATTTT 1712
QY 67 TAAATCTTATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 126
DB 1713 CATATTTATTAATCACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1772
QY 127 ATATTAATAAGTTAATAAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 186
DB 1773 CAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1832
QY 187 TTAACAACAAGA--GTAATAGTGAATCAACCCAAAGCAAGAAAGTTCTTGT 243
DB 1893 TAAATGAAGAAGAGGATAGTTTGCACCTTAACAAGTGAACAAGCAATTTGATGAT 1892
QY 244 ATATTAATCGGGCCGAGTGGCTGAAGAAGCCAAAGCAGAGGCTTAAAGCTCAACCAAC 303
DB 1893 AGTTGTACTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1952
QY 304 CGAAGCATTGTCTAATTAAGTCCATATTAAGCAAGAGGCGCGGGAAGAAAGAAAC 363
DB 1953 TGAAGCACTGTCTTAATCAAGATGAATTAATTAAGAGCGCGCTGATGG--TAAAC 2009
QY 364 CGTTGCCAGCTTATGAAGAGTGCATGCACTTTTGAAGAAAGTAAGTAAGTAAATCCCG 423
DB 2010 GGTAGCTGAATCATGAGCTATGAAGAAACAATTTTAAACGAGAAAGTGCATGATGG 2069
QY 424 GGTGGCTAATATGTTCCCGATCTAGGTGTGAAGCCACCTTCTGATGTGACAACT 483
DB 2070 CGTAGCTAATCATGATTAAGAACTGAAATGAAACCACTTTCCAGATGGTACAAAGT 2129
QY 484 TGTATCTGTGAATGGCCCATGAGACGATGAGCACTTCAAGAGGCG--GAAGTGAAT 542
DB 2130 AATAACAGTCCATCAACCAATGTTTAAAGAGCTATATCAATGATTCCTGTGAAATTA 2189
QY 543 TTGGTTGGCAATGAACATCGAGCTCAATGACAGCAAGAAAGTAACCGAATCTGAGTTA 602
DB 2190 TTGTTAAATAATCTGAATAGAGTCAATTAACATCATCCGAAACGGTGAATTTGAAGTA 2249
QY 603 CTATGAAGGCTTAATCTTGCATGTGGGTAGCCATTTCACTTTTGAAGCTAACA 662
DB 2250 AAAATACAGGCGATAGACTATACAGTAGTTCATTCCATTTTGAAGCAATA 2309
QY 663 AGGCACTAAATTTGATGTGTGAAGAAAGCTATGCAAGCCTATATATTTCCCTCTGGA 722
DB 2310 AAGCAATTAATTTATCTGTGAAGAAAGCAATATGTGAACATTTGATATTTCTGCGAG 2369
QY 723 ACACGCTACGATTTGGGCGAGCAAAACCGCAAGTGCATTTGATTCCTCTGTGGTGA 782

DB 2370 CTGAGTAGATTTGAACCTGAGATGAAAAAAAGTAACTTGTGCAATATCTGAGAC 2429
QY 783 GTAATAAAGTATGGCATGAACGGGCTTGTGAAT---ACATCGCGATGAACGCCATA 839
DB 2430 GACGTAAATTTATGATTTCCGTGTATGTGATGATGGGATATTTGACGAAGAACCGAT 2489
QY 840 AACATMAAGCGCTTGACAAAGCGAAATCTCAACGATTTATCAAGTAAGAGACTCCCATG 899
DB 2490 TCCGTCAATGATTCAAATCAAAAACCGCGCTTAAAAAACGATCAGCGCAAGCAATG 2549
QY 900 -----AAAAAGAAAAACAAGATATGTAA 925
DB 2550 CGAATMAAAGGTGTAAATTAAGCATGATTTTAAATGACACATCTCATATACAAAG 2609
QY 926 TACCTACGGAACCAAGCGATTAAGTGGCTTGAAGATATACGATCTTTGGGCGAGA 985
DB 2610 TCTTTATGACCACTGATGAGAGACTGTGTGAGATTTAGAGATACGAACCTTGTTCACA 2669
QY 1046 CCGTAGGGTATGGTCA-----GCAATAGCCCTGATGAAGAACCTTGATTT 1096
DB 2730 TCGTATGATGTGGCTCAAAATCTTAATGTGACAAAGATGATTAATAATGTACCGATTT 2789
QY 1097 AGTCACTAAGCGGATGATTAATGACATCAACGGGATTTAACAAGCGCATTTGGAT 1156
DB 2790 AGTTTAATCTAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2849
QY 1157 TAAAAAACGCAAAATCAATGATGATGCAAGGCAAGAAACAAGACATGCAAGATGGCGT 1216
DB 2850 TAAATAATGTTATATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 2909
QY 1217 AAGCCCTATATGTGTGGGTGGGCAAGAGCACTAGCAGGGAAGGTATGATTT 1276
DB 2910 -----TGACATCATCATTTGGTGGCAACATGATTTATTTGCTGCTGAAGTAAATGT 2963
QY 1277 TACCGCTGGGGAATGCAATTAACACCACTTCTTCTTCCACAAATTCCTACCGC 1336
DB 2964 TACGCGCGGATATGATACACAGTGCATCTCAATCTGAAACAAGCTGAAGTTGC 3023
QY 1337 TCTAGCAATGCGTGAACCATGATTTTGAAGCGCGCAGAGTCTGTATGATGACGA 1396
DB 3024 ACTTAGAGTGTATTAACAACCATATGCGTGAAGAACTGGTCTTCTGAAGTGTCTA 3083
QY 1397 TCGCATCTACTCTCGGCGAAATGAACCTTGAACCGCATGTTGCCGACAGAGA 1456
DB 3084 AGCGACTACTGTAACACCAAGCACTTGGCATTTTCATGCGCATGTTAAGAGCAGAGAGA 3143
QY 1457 GTATTTCTATGAATGTGGGCTTTTGGCAAGGCAATAGCTCTAGCAAAAACAATTGT 1516
DB 3144 GATGCTTAATTAATGAGATTTATCTGTAAAGTCAAGCTGCAATCATATCTGCACTAT 3203
QY 1517 AGAACAAGTGAAGGCGGCGATGTTTAAATTTGATGAAGACTGCGGCGACAACACC 1576
DB 3204 TGAACAAATTCATGAGGCGCTATAGCTTTAAATGATCAATGAAGATGGGAGCTACAC 3263
QY 1577 AAGTCGATCATCACTGCTTGAAGCGTGCAGATGAATATGATGTGAAGTTTATCA 1636
DB 3264 TTCACGATTAAGTCAATGATTAAGCGTTGACATGATGATGATGATGATGATGATGAT 3323
QY 1637 CACCGATACATGATGAGGAGGTTATGTATGATGACACCTTAATGCAATGAACGGGG 1696
DB 3324 TGCAGATATTAATTAATGAAGCTGATTTATGAAGATTAACAAAGCTGCTGTAAAGATG 3383
QY 1697 CGCATTCATGCTCAACATGAGGAGCGGTTGAGAGCACTCATGATTTATCA 1756
DB 3384 TGTATTTGATATGATCATATGAGAGAGCTGTGTGTATGACCTGATCTATATCA 3443
QY 1757 CATGCAAGCGCATCAATATTTCAACCTCT 1788
DB 3444 ATCAGTGCATATTAACATCTTACTTCT 3475

```
RESULT 6
US-10-834-151-2/c
; Sequence 2, Application US/10834151
; Publication No. US20050244836A1
; GENERAL INFORMATION:
; APPLICANT: TSANG, TAT-KIN
; APPLICANT: MENG, XIANGWEN
; APPLICANT: ZHANG, HONGJUN
; TITLE OF INVENTION: METHODS AND COMPOSITIONS TO DETECT BACTERIA USING MULTIPLEX PCR
; FILE REFERENCE: 37611-95516
; CURRENT APPLICATION NUMBER: US/10/834.151
; CURRENT FILING DATE: 2004-04-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 2
; LENGTH: 526
; TYPE: DNA
; ORGANISM: Helicobacter pylori
US-10-834-151-2

Query Match      7.5%; Score 216.2; DB 1; Length 526;
Best Local Similarity 70.9%; Pred. No. 1.7e-43;
Matches 287; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 900 AAAATGAAAAAACAAGAAATATGTAAATACCTACCGACCCACAAAGGCGGATAAAAGTGGCG 959
DB |||||
QY 960 TTAGGAGATACCGATCTTTGGGCGAGAAGTGAACATGACTATACCACTATGGCGAAGAA 1019
DB |||||
QY 345 TTGGCGGATACAGACTTGATCGCTGAAGTGAACATGACTACCACTTATTTGGCGAAGAG 286
QY 1020 CTTAAATTTGGCGGGTAAACTATCGTGAGGGTATGGTTCAGAGCAATAGCCCTGAT 1079
DB |||||
QY 285 CTTAAATTTGGTGGCGGTAACACCTGAGAGAAGGCATGAGCCAATCCAAACCCCTAGC 226
QY 1080 GAAAAACCCCTAGATTTAGTCACTACCAACGCGATGATTTCGACTACACCGGGATTTAC 1139
DB |||||
QY 225 AAAGAAGATTTGGATCTAATCATCACTAACCGCTTTAATCGTGGAATTACCCGGTATTTAT 166
QY 1140 AAAGCCGACATTTGGGATTTAAAAACGCGAAATCCATGGCATTTGGCAAGCGGCAAAACAAG 1199
DB |||||
QY 165 AAAGCGATATTGTTAATGAAGTGGCAAAATCGCTGGCATTTGGTAAAGCGGTAAACAA 106
QY 1200 GACATGCAAGATCGCGTAAGCCCTCATATGCTCGTGGGTGGGCGACAGAAGCACTAGCA 1259
DB |||||
QY 105 GACATGCAAGATCGCGTTAAAAACAATCTTAGCGTAGGTCCTGCTACTGAAAGCCTTAGCC 46
QY 1260 GGGGAGGTATGATTATACCGCTGGGGGATCGATTTCACACACC 1304
DB |||||
QY 45 GGTGAAGGTTTGTATCGTAACTGCTGGTGGTATTGACACACATC 1

RESULT 7
US-10-793-626-121/c
; Sequence 121, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793.626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 121
; LENGTH: 309
; TYPE: DNA
; ORGANISM: Artificial Sequence

US-10-793-626-121/c
; Sequence 121, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793.626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 121
; LENGTH: 309
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-121

Query Match      3.4%; Score 97.4; DB 1; Length 309;
Best Local Similarity 59.4%; Pred. No. 7.9e-15;
Matches 186; Conservative 0; Mismatches 121; Indels 6; Gaps 1;

QY 1114 TGATTATCGACTACACCGGGATTTACAAAGCCGACATTTGGGATTTAAAAACGCGCAAAATCC 1173
DB |||||
QY 308 TAAATTATTGATTATGACAAGATTGTTAAAGCAGATATCGAATTAAAAATGGTTATATTT 249
QY 1174 ATGGCATTTGGCAAGCGCAGGAAACAGGACATGCAAGATGGCGTAAGCCCTCATATGGTCG 1233
DB |||||
QY 248 TTAAGATCGGTAAAGCTGGAACCCAGGATATAATGATAACGT-----TGACATCATCA 195
QY 1234 TGGGTGTCGGCACAGAACGACTAGCAGGGGAAGGTATGATTATTACCGTGGGGGATCG 1293
DB |||||
QY 194 TTGGTGCAACAACCTGATATTATTGCTGCTGAAGGTAAAAATTGTTACTGCCGCGGTATCG 135
QY 1294 ATTACACACACCCACTTCTCTTCTCCACAACAATTCCTACCGCTCTAGCCAATGGCGTTA 1353
DB |||||
QY 134 ATACACACGTCGACTTTCATCAATCTGAACAAGCTGAAGTTGCATTTGAGAGTGGTATTA 75
QY 1354 CAACCATGTTTGGAGCGGCACAGTCTCTGTAGATGGCAGCAATCGCACTACTATCACTC 1413
DB |||||
QY 74 CAACGCATATCGGTGGAGGAACTGCTGCTTCTGAAGGTGCTAAAGCGACTACTGTAAACAC 15
QY 1414 CGGCAAAATGGAA 1426
DB |||||
QY 14 CAGGACCTTGGCA 2

RESULT 8
US-10-793-626-115
; Sequence 115, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793.626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 115
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Artificial Sequence

US-10-793-626-115
; Feature:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence

US-10-793-626-115

Query Match      2.6%; Score 74.2; DB 1; Length 336;
Best Local Similarity 56.9%; Pred. No. 3.5e-09;
Matches 136; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 594 TTGAGGTTACTTAATCAAGGCGCTAAATCCTTGCATGTGGTAGCCATTTCCACTTCTTTG 653
DB |||||
QY 5 TTGAAGTAAAAAATACAGCGCATAGACCTATACAAGTAGGTTCACTTTCCACTTTTCG 64
QY 654 AAGCTTACAAGGCACTAAAAATTCGATCGTAAAAAGCCCTATGGCAACGCCCTAGATATTC 713
DB |||||
QY 65 AAGCAATTAAGCATTTAGAAATTTGATCGTGAAGAGCATATGGTAAACATATTTGGATATTC 124
QY 714 CCTCTGGCAACACGCTAGCGATTTGGGGCAGCAAAACCCGCAAGTGCAGTTGATTCCTC 773
DB |||||
QY 125 CTGCAGGAGCTGCAGTGAGATTGAACTGGAGATGAAAAAAGTACAACCTTGTGCAAT 184
```


[illegible]

```

; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4033
; LENGTH: 3591
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4033

Query Match          1.3%; Score 37.8; DB 1; Length 3591;
Best Local Similarity 53.3%; Pred. No. 7.5;
Matches 72; Conservative 3; Mismatches 60; Indels 0; Gaps 0;

Qy 23 AGCACATATTGATCCGTGTGTGGTGGTGGTAAATTCRACCTCTGTTAATRCCTATTATTAAT 82
Db 2984 AGCTCGATGGAGTTTTTTAGTACCTCGAAATTTTACAACGTAGATGAGCTACTATTCAA 2925
Qy 83 TTTTAAATAATTACTTTATTCATATATATAATAATTAATTAATTAATTAATTAATTAATTAAT 142
Db 2924 TCITTTATCATGTTCTAATGATGATTTTAAACGATATGAATTCGATTGTGAAATGATTAAT 2865
Qy 143 AAAAAGTAACGAAAT 157
Db 2864 GAGAAACGTCACAT 2850

RESULT 15
US-10-793-626-4132/c
; Sequence 4132, Application US/10793626
; Publication No. US2005025478A1
; GENERAL INFORMATION:
; APPLICANT: KIMWELLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUB480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4132
; LENGTH: 2988
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4132

Query Match          1.3%; Score 37.2; DB 1; Length 2988;
Best Local Similarity 52.7%; Pred. No. 9.5;
Matches 78; Conservative 1; Mismatches 69; Indels 0; Gaps 0;

Qy 64 TGTTAATRCCTATTATAATTTTTTAAATAATTAATCTTATTCATATATATAATAATTAATTA 123
Db 509 TTTTATATGTTAAAGATTTCTTTTTTAAACCTAAATAAATTTTATTAACCTGTGTTATTA 450
Qy 124 CTTTATATTAAGAGTTTAAATAAAGTAACGAAATTAGGACTATAATCCCATGTCCTTTAA 183
Db 449 CTAAATAATCAATTTTATACATCCCTTAAAGTCTATTTATTTTGTCATTTTGACTAATA 390
Qy 184 AATTTAACACAAGGAGTAAATAGTGAAA 211
Db 389 TACTTGTCAATAGACAATATATTATA 362

Search completed: November 29, 2005, 08:13:49
Job time : 569 secs

```

Search completed: November 29, 2005, 08:13:49
Job time : 569 secs